THEORY OF OPERATION
AND
APPLICATIONS OF SAMPLED ANALOG DEVICES
IN RECURSIVE COMB FILTERS

Stacy Vernon Holmes

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

THEORY OF OPERATION
AND
APPLICATIONS OF SAMPLED ANALOG DEVICES
IN RECURSIVE COMB FILTERS

by

Stacy Vernon Holmes

June 1976

Thesis Advisor: T. Tao

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Theory of Operation and Application of Sampled Analog Devices in Recursive Comb Filters		5. TYPE OF REPORT & PERIOD COVERED Ph.D. Thesis; June 1976
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Stacy Vernon Holmes		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940.		12. REPORT DATE June 1976
		13. NUMBER OF PAGES 332
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Charge-Coupled Devices
Comb Filters
Sampled Analog Filters
Moving Target Indicator Simulation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The theory of operation of charge coupled devices in the sampled analog mode is investigated. The applications of charge coupled devices in recursive filters is explored, both in first and second order systems . A means for translating nth order filters to first

DD FORM 1 JAN 73 1473 (Page 1) EDITION OF 1 NOV 68 IS OBSOLETE S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(20. ABSTRACT Continued)

or second order is explained. Tables are developed to assist the engineer in choosing the proper coefficients for second order recursive filter design. Applications of recursive filter systems to moving target indicator processors are explored.

Theory of Operation
and
Applications of Sampled Analog Devices
in Recursive Comb Filters

by

Stacy Vernon Holmes
Lieutenant Commander, United States Navy
B. A., Northwestern University, 1961
E. E., Naval Postgraduate School, 1968

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

from the

NAVAL POSTGRADUATE SCHOOL
June 1976

ABSTRACT

The theory of operation of charge coupled devices in the sampled analog mode is investigated. The applications of charge coupled devices in recursive filters is explored, both in first and second order systems. A means for translating nth order filters to first or second order is explained. Tables are developed to assist the engineer in choosing the proper coefficients for second order recursive filter design. Applications of recursive filter systems to moving target indicator processors are explored.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

The author expresses his greatest and sincerest appreciation to Dr. T. F. Tao for his guidance, tolerance, insight, assistance, and encouragement during the time of this study. Without such assistance, the thesis could not have been completed.

The author expresses his gratitude also to the other members of the committee, Dr. D. E. Kirk, Dr. D. R. Barr, Dr. P. C. C. Wang, and Dr. R. Panholzer for their interest and assistance.

The author is also indebted to many members of the solid state research team. Members who have provided invaluable assistance in the experimental part of the research are ENS Bruce Freund, USN, LCDR Lars Saetre, Norwegian Navy, FLT/LT Arif Ejaz, Pakistan Air Force, LT Jim Campbell, USN, LT Frank Piazza, USN, LT Ang Vong Mongkol, Cambodian Navy, and CDR Kurt Pfennig, Federal German Navy.

Special thanks go to first class petty officer Phil Kopp whose ever ready technical assistance was invaluable. Thanks also go to Alicia, Albert, and Alfred Tao for their computer work.

INTRODUCTION

It was the intent of this research to consider the potential signal processing use of the charge coupled and related devices and how their parameters might help or hinder such applications. In order to retain this overall perspective, the research was pursued on three levels concurrently. Those levels are

A. Charge coupled device physics . . . the physical explanation of the operation of the device in engineering terms. We mean to emphasize the parameters the engineer is most interested in. A list of such parameters include gain, linear range, dynamic range, bandwidth, and maximum delay. The determination of the relationship between these parameters and the voltages which could be used to clock, bias, or input/output the device was paramount.

B. The CCD as a circuit element. The object of this level of the research was to determine how the CCD functioned as a sampled-analog delay device. The natural application was in delay line filters and both first and second order devices were built and tested, relying heavily on the characterization performed in A. and also on our growing knowledge of filter design, theory and practice. Particularly emphasized in our research was the recursive comb filter since it appeared that in this area we might make a significant contribution to technology.

C. The application of the circuit element (the recursive comb filter) to major electronic systems. System specifications were determined to have a profound effect on the devices selected for implementation. It was our goal to seek out these specifications and adapt our CCD's (or, in some cases charge transfer devices) to the system at hand. The most widely known application of the periodic frequency transfer function is the radar system moving target indicator (MTI). Physical realities were assessed to evaluate the applicability of the CCD filter.

We believe that this purpose has been, for the most part, achieved. Chapter one is the treatment of the parameters of the CCD. Chapter two follows with an exhaustive treatment of the filter design problem including the b-plane triangle of stability derivation. Finally, chapter three deals with the radar system and how requirements dictate, or at least, suggest, use of various filter designs.

Each of the chapters is capable of standing alone. The connecting link among them is the use of a sampled-analog device as the primary circuit element.

CHAPTER ONE

THE CHARGE COUPLED DEVICE AND ITS ENGINEERING CHARACTERISTICS

1.0 Introduction.

Boyle and Smith's paper [1] in 1970 introduced the concept of causing packets of charge at the surface of a semiconductor to move about on the surface by moving the potential minimum (or maximum). Their brief paper opened the door to a new technology which has grown so quickly that already three international conferences [73,74,75] on the applications of charge coupled devices have been held, each opening new frontiers of applications of these devices. The reasons for such wide acceptance can be found partially in their unique two-dimensional characteristics which permit the simple design of imaging devices. This application was perceived early, and both theory and experimental results were described as early as 1971 [2,3].

Another of the attractive characteristics is the simplicity of fabrication, which rendered the CCD easier to manufacture than the older BBD (bucket brigade devices) [4].

Finally, the charge coupled device is capable of being used as a sampled analog circuit device. Unlike most recent semiconductor developments which are digital in nature, the CCD can and has been used in analog signal processing. In particular, the CCD will almost surely replace the older

forms of analog signal delay lines (quartz crystals, lumped circuit parameters) in many systems since it is

1. Smaller
2. Requires no tight mechanical specifications
3. Potentially cheaper
4. Less sensitive to temperature variations
5. Capable of providing longer delay (=1-5 seconds)

The CCD was first used as the delay element in a transversal filter in 1972 [5]. Recursive filter design using CCD's was not reported until 1974 [6] and by our own research at NPS in 1975 [7].

Section 1.1 The Charge Coupled Device

The CCD is an array of MOS (metal-oxide-semiconductor) capacitors, each with a separate electrode. These electrodes are pulsed in order to move a packet of charge about on the substrate. This packet bears a relationship to the information which was used as the input. In case of digital input, the packet is present, or not. In the more interesting case (for our use) of analog input, it is hoped that the size of the packet is related, preferably linearly, to the input.

First we wish to look at the structure of a typical CCD, shown in Figure 1.1. Included are the metal electrodes, the dielectric oxide, and the n-doped silicon. This implies a p-channel device. The ϕ's are the clock pulses. Notice that in the configuration in Figure 1.1a, the oxide thicknesses

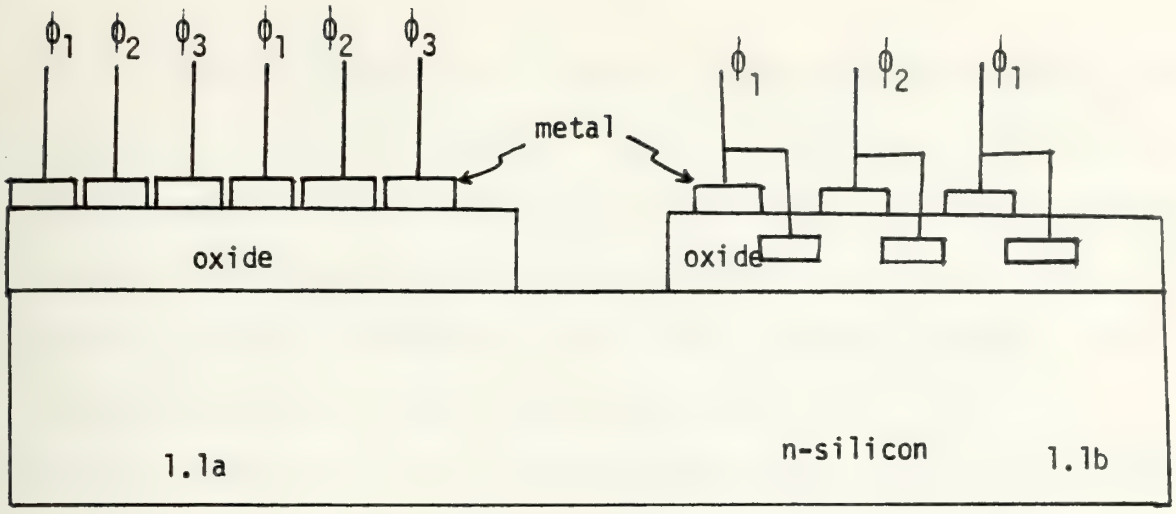


Figure 1.1 CCD structure

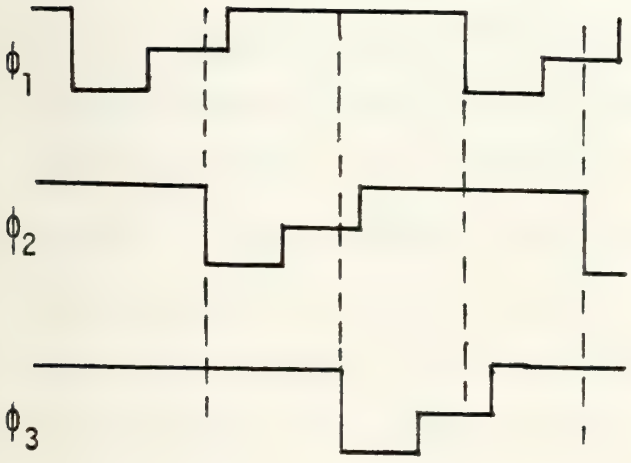


Figure 1.2a

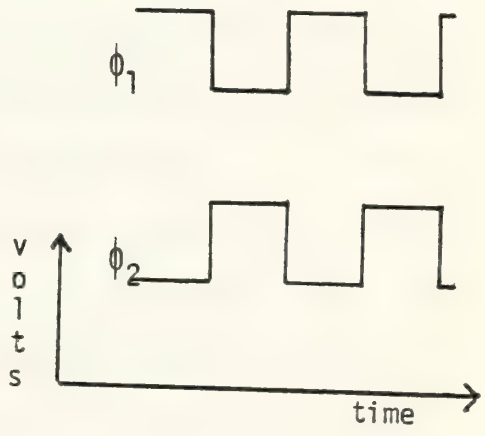


Figure 1.2b

Figure 1.2 Pulse trains necessary for proper operation of the CCD. Figure 1.2a shows THREE PHASE operation, Figure 1.2b shows TWO PHASE operation.

under each electrode are the same, and that there are three separate clock pulse trains (shown in Figure 1.2a). On the other hand, the configuration in Figure 1.1b shows differing oxide thicknesses under electrodes and only two clock pulse trains (shown in Figure 1.2b). For obvious reasons, the former is called THREE PHASE operation, the latter TWO PHASE operation. The TWO PHASE operation, though more complicated in fabrication, is widely used because of the simplicity of clock pulse trains which can drive it. Figure 1.2 shows each pulse train. In the THREE PHASE operation, three distinct levels of voltage are observed. If three adjacent electrodes are considered, it becomes clear that transfer is effected at the times indicated by a dotted line. At the first dotted line, transfer of the charge stored under electrode no. 1 to electrode no. 2 is effected. At dotted line no. 2, transfer is made from 2 to 3, and so forth. This requires the three voltage levels to insure that transfer is always made in the correct direction. As a matter of fact, this points up an advantage of three phase operation in that it can be used to transfer charge in any direction and depends only on the magnitude of voltages used. As we will see, in TWO PHASE operation, the direction of transfer is determined by the structure of the oxide thicknesses.

In the TWO PHASE operation, we see the transfer of charge being made when one pulse goes high and the other low. The

differences in the oxide thickness enable us to use a two level clock pulse because the narrower oxide provides for more of the voltage to be dropped IN THE SEMICONDUCTOR instead of across the oxide. Thus the surface potential (silicon surface) under the two insulator thicknesses are different. The surface potential profile is given by Figure 1.3. Although this structure does not easily lend itself to backward transfer, it is widely used because of the simplicity of the clock pulse train.

We have stated that the CCD is an array of MOS capacitors. Some of the fundamentals of MOS capacitor characteristics are defined here. First of all, if the voltage across the MOS device is such to attract majority carriers to the oxide-silicon surface, then we say that we are operating the device in the accumulation mode. Figure 1.4 shows, at the top, the accumulation mode.

If, instead, we are operating the device in the reverse biased mode, so that the majority carriers in the silicon are repelled from the oxide-silicon interface, then we say we are operating in either the DEPLETION or the INVERSION mode. If the charge (positive) gathered at the interface is made up of positive ions, which are relatively immobile, then the DEPLETION mode is what we have. If holes appear as the positive charged particles, then we have achieved INVERSION ... that is, the carriers in the silicon are no longer electrons. Instead, the silicon has been inverted and the carriers are holes.

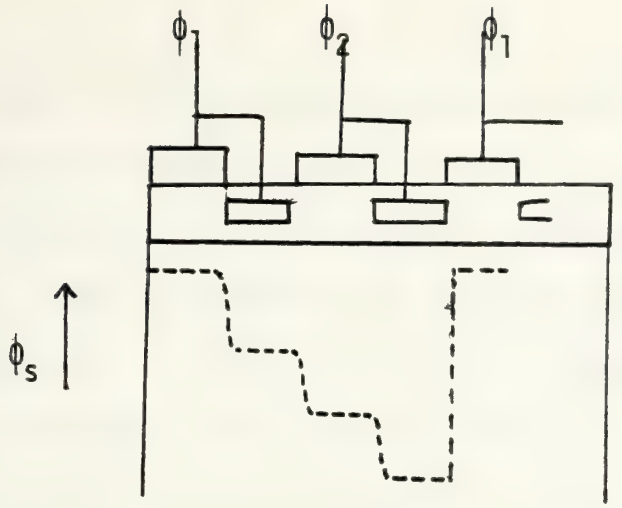


Figure 1.3 Surface potential, ϕ_s, profile of a TWO PHASE CCD.

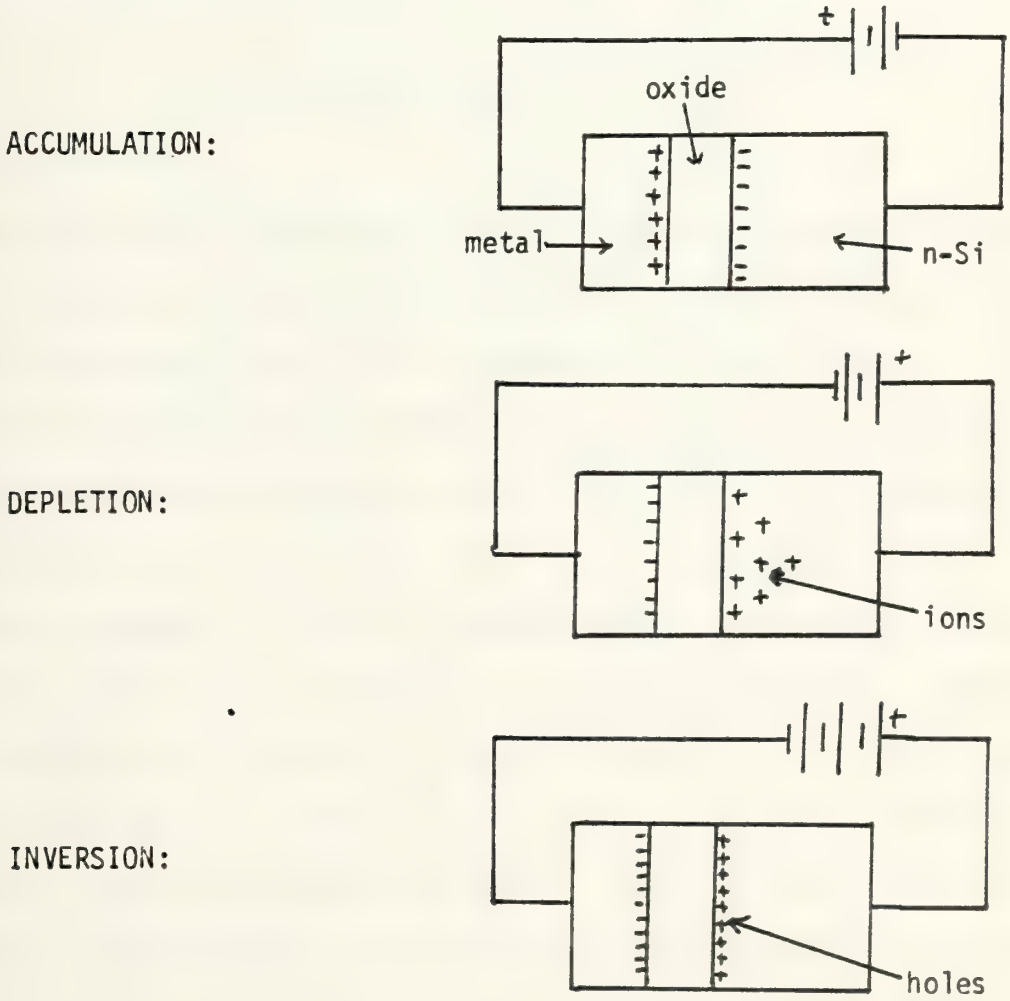


Figure 1.4 MOS capacitor fundamentals

Figure 1.5 shows the charge profiles and electric fields in the three cases. Notice that the electric field extends much further into the silicon for DEPLETION than for accumulation. This is because the charged region in the silicon is much wider in depletion than in accumulation. As inversion is reached, the electric field may one again extend only a brief distance into the silicon. Thus we note that the electric field is very dependent on the mode of capacitance we are utilizing.

Capacitance is defined by

$$C = \frac{\Delta Q}{\Delta V} = \frac{\varepsilon A}{d} \qquad (1\text{-}1)$$

where the latter equality is valid for parallel plate capacitors. Figure 1.6 shows the incremental charge accumulating at the interface due to an incremental voltage change. Notice that the distance between the charged regions is the same after the application of ΔV as it was before the application. This implies that the "d" in the above equation (1-1) remains constant. Thus the capacitance is constant. Now consider the case in depletion where the ΔQ forming because of the application of ΔV is made up mostly of positive ions. Here there are two different "d"'s no matter what convenient method is used for measuring them. Thus we must conclude that the capacitance is NOT constant, but that it depends on the applied voltage as well as the accumulated charge at the interface. Figure 1.7 indicates the variation of the

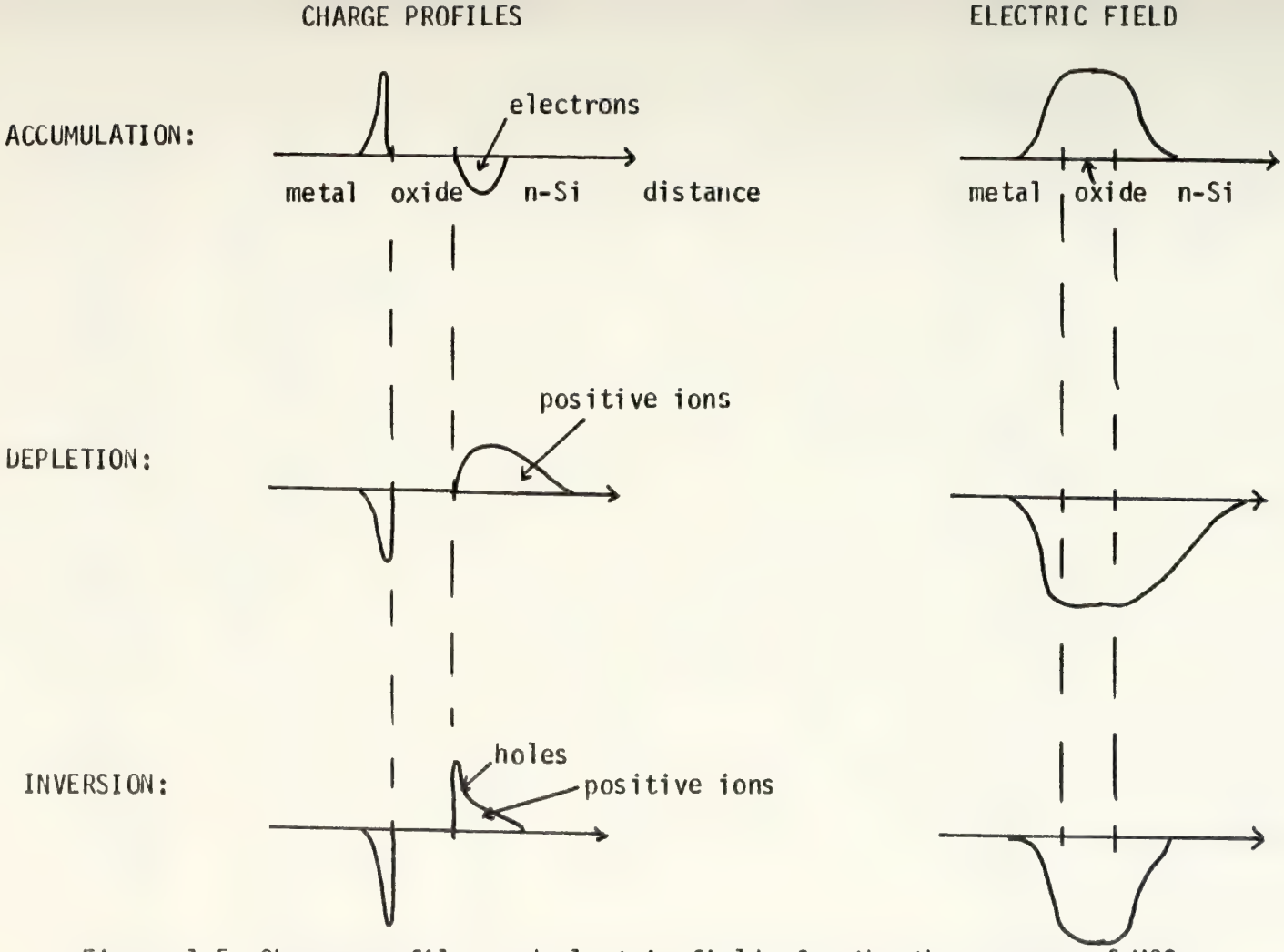


Figure 1.5 Charge profiles and electric fields for the three cases of MOS capacitance.

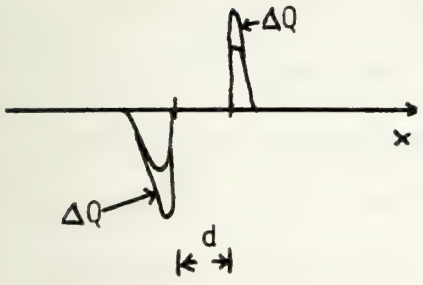


Figure 1.6a

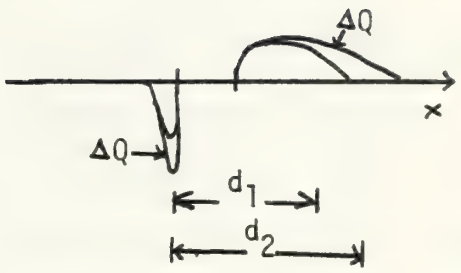


Figure 1.6b

Figure 1.6 Incremental charge accumulating at the interfaces of a MOS capacitor due to an applied ΔV. Figure 1.6a shows no change in d, while figure 1.6b shows a change in d. Thus, in the capacitor in 1.6b, capacitance is a function of applied voltage.

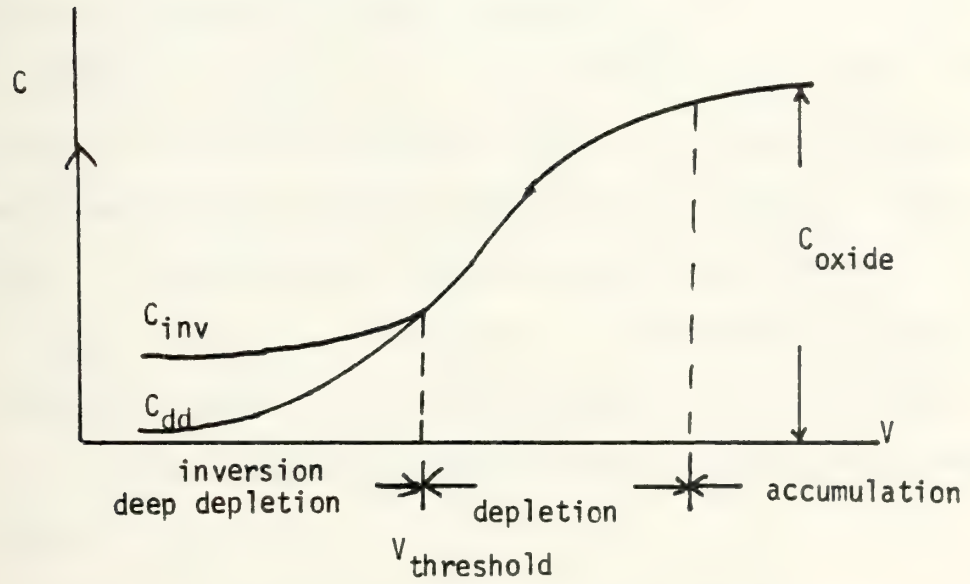


Figure 1.7 Variation of capacitance in a MOS capacitor. For our p-channel device, $V_{threshold}$ is a negative voltage. The ordinate axis appears at an arbitrary voltage.

capacitance. Notice that the capacitance is greatest when operating in the accumulation mode, and is least when driving the capacitor into what is called DEEP DEPLETION. Deep depletion occurs when a negative voltage exceeding the threshold voltage is impressed on the electrode. For low frequency operation, inversion, caused by the thermal generation of electrons and holes in the depletion region (dark current) is used. But if the MOS capacitor is pulsed very fast, the dark current cannot fill the potential "well" and thus the capacitance seen is that of deep depletion. This is the region in which the CCD operates. Thus we can see that the CCD, when used for storage of data, is a very volatile memory ... the absence of a clock pulse train renders the device quite useless.

Figure 1.8 shows the basic model of the CCD. Notice that it is two capacitors in series. The following are equations used to compute the actual values of this capacitance. The depletion capacitance is a function of the dielectric constants of both the silicon and the oxide, the voltage impressed on the electrode, and the doping level of the silicon, the oxide capacitance, and the thickness of the oxide.

$$C_{dep} = \frac{C_{acc}}{1 + \dfrac{2\kappa^2_{SiO_2}\ \varepsilon_o\ V_G}{q\ N_D\ \kappa_{Si}\ th^2_{SiO_2}}} \qquad (1\text{-}2)$$

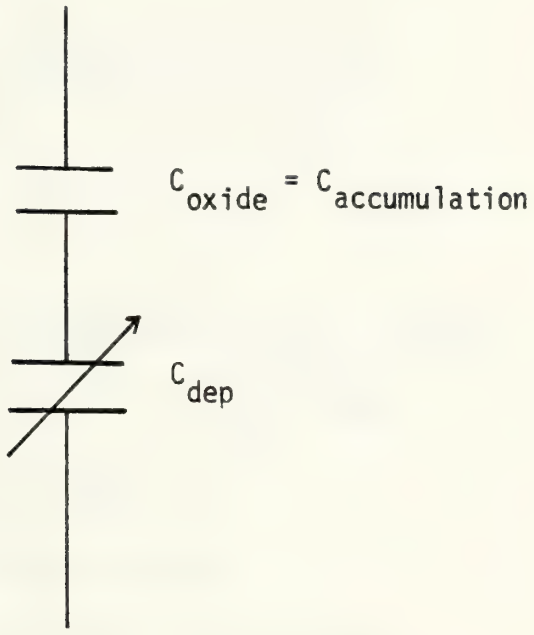


Figure 1.8 Basic model of a CCD: Two capacitors in series.

and

$$V_G - V_{FB} = \phi_s + \frac{1}{C_{acc}} \sqrt{2\ \kappa_{Si}\ \varepsilon_o\ q\ N_D\ \phi_s} \qquad (1\text{-}3)$$

where

κ_{SiO_2} = relative permittivity of Silicon Dioxide

ε_o = permittivity of free space

V_G = gate bias voltage

q = charge on electron

N_D = doping density of silicon

κ_{Si} = relative permittivity of silicon

th = oxide thickness

V_{FB} = flat band voltage

ϕ_s = surface potential of silicon

Equation (1-3) indicates that the equations can be broken down even further in that the really important parameter is not the voltage on the electrode, but the surface potential of the silicon. Equation (1-3) gives the relation between the gate voltage and the surface potential. These equations appear in many references [1] [4].

Some useful definitions appear below.

$$\tau_{st} = \frac{Q_{st}}{I_{dark}} \qquad (1\text{-}4)$$

$$\tau_o = L^2/4D \qquad (1\text{-}5)$$

$$\eta = \Delta Q/Q; \qquad \varepsilon = 1 - \eta \qquad (1\text{-}6)$$

Storage time, τ_{st}, is defined as the time it takes to completely fill the potential well solely by dark current ... that is, by the thermal generation of holes and electrons. This number is on the order of seconds which provides us with a good bandwidth of frequences which will permit proper operation. The full well charge is defined as

$$Q_{st} = V_G (C_{inv} - C_{dd}) \qquad (1\text{-}7)$$

where C_{inv} is the capacitance at inversion and C_{dd} is the capacitance at deep depletion.

The second parameter is transfer time. L is the length of the electrode and D is the diffusion constant of the material. Actually, transfer is complicated and one time constant is not enough to characterize completely the transfer operation. The first CCD paper ever published, Boyle and Smith [1], indicated that the transfer time constant was equal to the length of the electrode squared divided by four times the diffusion constant. Subsequent researchers have found that transfer is actually a combination

of three different transfer mechanisms and each has a different time constant. Also, their time constants change with time and distance, so the transfer time is quite complicated. We shall look at some of the basics of this subject later.

The third parameter is transfer efficiency, or analogously, transfer inefficiency. Inefficiency is defined as the fraction of charge left behind during each transfer. Since devices are made which get an efficiency of 99.99%, it is a good approximation to say

$$Q_n \approx Q_o(1 - n\varepsilon) \qquad (1\text{-}8)$$

where Q_o is the magnitude of charge in the zeroth transfer gate and Q_n is the magnitude of charge in the nth gate.

If this inefficiency is larger than normal, we may have encountered severe noise problems. The main noise problem is in the interface states of the silicon-oxide boundary. These states trap holes and hold them long enough so that they are emitted in a subsequent charge packet. Other sources of noise are listed below:

A. "Edge Effects"

B. Bulk Traps

C. Dark Current

D. Thermal Noise in the input/output circuits

The presence of the interface states have led to a method of alleviating this source of noise (to first order) by simply insuring that these trapping states are filled at all times. We do this by what is called the FAT ZERO. In order to accomplish the filling of the states at all times, we must use, for n-silicon (a p-channel device), a voltage for the HIGH phase of the clock which is near the threshold voltage. As we can see, the introduction of charge for all time in a potential well means that the "well" is never totally empty. Nevertheless, we adjust our thinking to the idea that zero is actually some finite size of charge packet. Since the trapping states are always filled, the signal is passed without distortion (to first order).

Section 1.2 Input/Output Considerations

We will now look at some of the input schemes of the CCD. We wish to operate the CCD in the deep depletion region, but introduce minority carriers (holes) to the electrodes in accordance with some desired signal. The first method of input is called DIODE CUTOFF and the physical arrangement is shown in Figure 1.9 The signal voltage is added to the bias voltage on the input diode. The diode is forward biased by this combination of bias and signal. The first input gate is pulsed, thus sampling the signal voltage. When the gate voltage is high, no charge is injected into the n-silicon region. But when the gate is driven low (the same level each time), charge is injected which is related

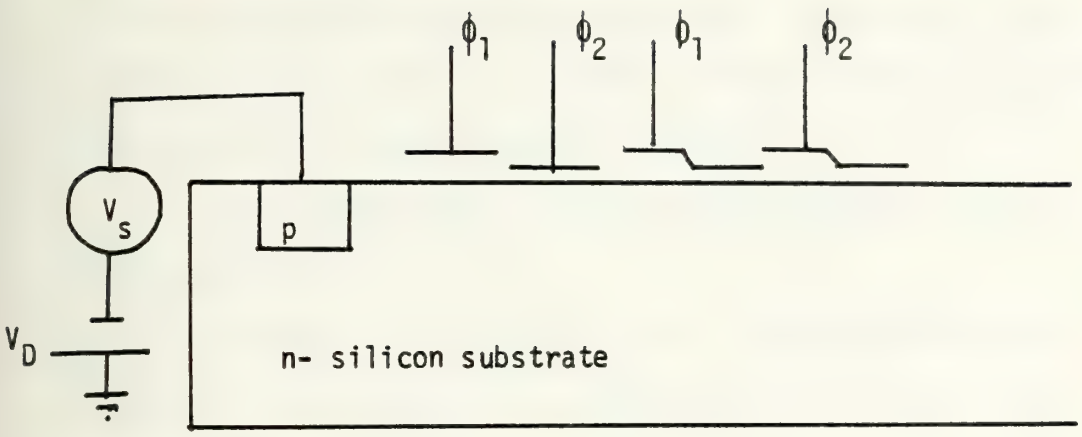


Figure 1.9 DIODE CUTOFF method of input

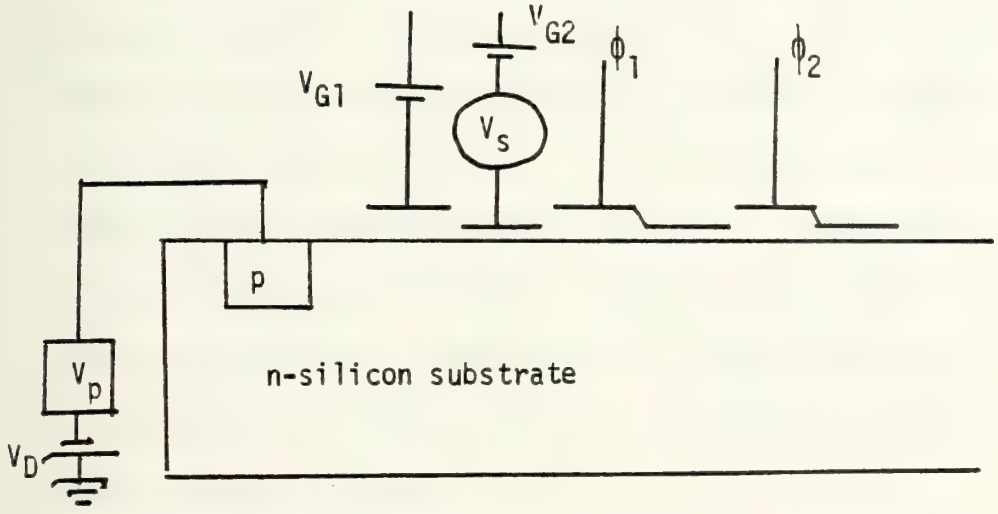


Figure 1.10 "Scuppering", "Sloshing" or Potential Equilibration mode of input to CCD. V_s is the signal V_{G1} acts as a constant parametric voltage, and V_p is the diode pulse.

to the signal voltage in a squared manner. That is, the size of the charge packet injected is proportional to the square of the signal voltage. Thus we have two disadvantages to the diode cutoff method of input:

1. The pulsed gate voltage must always reach the same voltage level.

2. We begin our operation in a nonlinear manner . . . that is, the very first step toward using the CCD, input, is nonlinear.

For these reasons, we believe that the diode cutoff method is not an attractive input mode.

The BARRIER mode is much like the diode cutoff mode. Here the signal "rides" atop a pulse which is applied to the diode. The gate voltage is held low. At the arrival of a diode pulse, a certain amount of charge is injected into the first potential well. It is subsequently transferred into the next transfer gate and so on. However, here too, we must have precise pulse requirements, and even then, the injected charge is theoretically proportional to the square of the signal voltage. This nonlinearity makes this mode unattractive as well.

The third mode of input is called POTENTIAL EQUILIBRATION, or "SLOSHING", or "SCUPPERING". Here the diode pulse is used simply as a source of holes to fill the well/wells while the signal voltage is applied to one of the input gates. At the end of the diode pulse, the excess charge under the

input gates is drained back across the reverse biased p-n junction at the input diode. The residue charge is linearly related to the signal voltage at the time of the diode pulse. This is best seen by considering Figures 1.10 1.11 and 1.12. Four instants in time are shown as indicated in Figure 1.11. The first time, t_o, occurs prior to the arrival of the diode pulse. The surface potential profile is drawn under the electrodes in Figure 1.12. Time t_1 is during the diode pulse. Holes have been injected by the forward biased diode into the wells under both G_1 and G_2. Although the holes accumulate at the surface, it is possible (and very illuminating) to draw them as if they fill UP the potential well. After all, when a hole arrives in a region, the potential of that region is increased marginally. The third time, t_2, occurs after the end of the diode pulse. Since the p-n junction is reverse biased, and since the surface potential of the silicon near the junction is very negative, the charge drains across the p-n junction leaving only the charge which cannot "climb" the potential barrier posed by the constant G_1 voltage. We have, thus, isolated a packet of charge under G_2 which is linearly related to the input signal voltage. Time t_3 shows the arrangement after the low value of ϕ_1 arrives. The charge is transferred from G_2 to the lower level of the first transfer gate. For clarity, the steps of "scuppering" are listed here in order:

A. Impress signal voltage onto G_1 or G_2

B. Bias G_1 and G_2 in accordance with operating curves

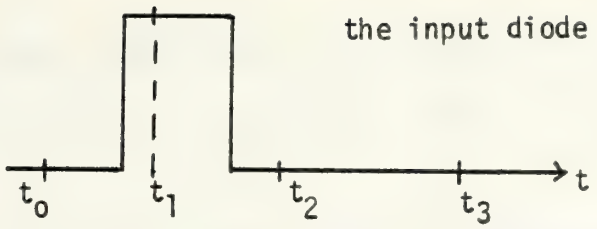


Figure 1.11 Time instants relative to the ON/OFF cycling of the diode input pulse.

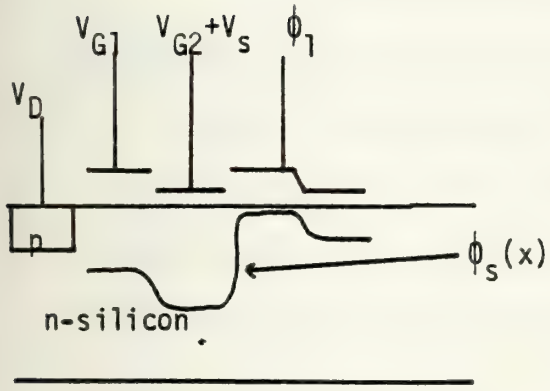


Figure 1.12a. $t=t_o$. Both gates 1 and 2 are empty.

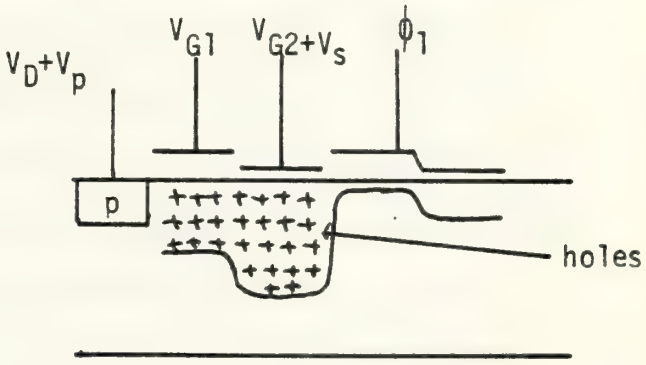


Figure 1.12b. $t=t_1$. Injection of holes fills gates 1 and 2.

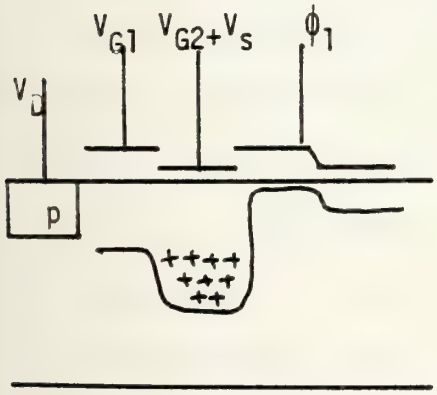


Figure 1.12c. $t=t_2$. Excess holes drain back across p-n junction

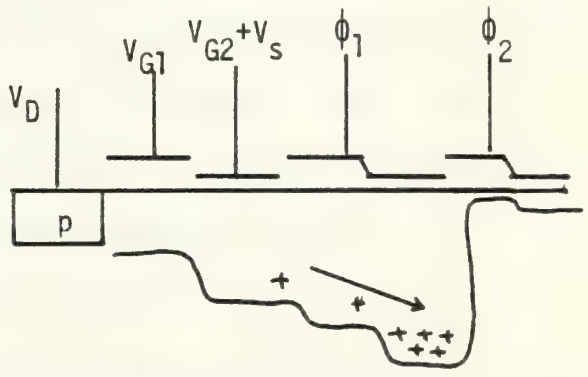


Figure 1.12d. $t=t_3$. Holes transferred to first transfer gate when ϕ_1 goes low

Figure 1.12. Sequential action of the CCD in the scuppering input mode. Shown is the stored charge, the input gates, and the surface potential.

C. Pulse the diode "ON" briefly

D. Inject holes into the potential wells under G_1 and G_2

E. At end of diode pulse, the holes will move back across the p-n diode junction leaving a portion of the injected charge.

F. The magnitude of the residual charge packet is LINEARLY related to the magnitude of the signal voltage.

Figure 1.13 shows the output stage. Notice that there exists an output gate designated by V_o. This V_o is generally lower than any of the clock voltage levels which facilitates movement of charge from the last transfer gate, via the output gate, and then across the reverse biased junction of the output diode. This charge (holes) collects on the capacitance at the gate of the output FET, thus turning the FET off ... or at least, reducing the current through it. At this time, the REFRESH FET is OFF, so that the charge simply stays there. When the refresh pulse goes very low, this charge is quickly drained off through the refresh FET to V_{DD}. After draining off, the gate capacitance of the output FET is empty of charge and now just awaits the next output pulse. The timing of these voltage levels is shown in Figure 1.14.

Figures 1.15 through 1.18 show the input-output characteristics achieved in our laboratory with a TRW, eight delay stage, four input device, by the scuppering input method. The segment of the curve on each figure toward the left is

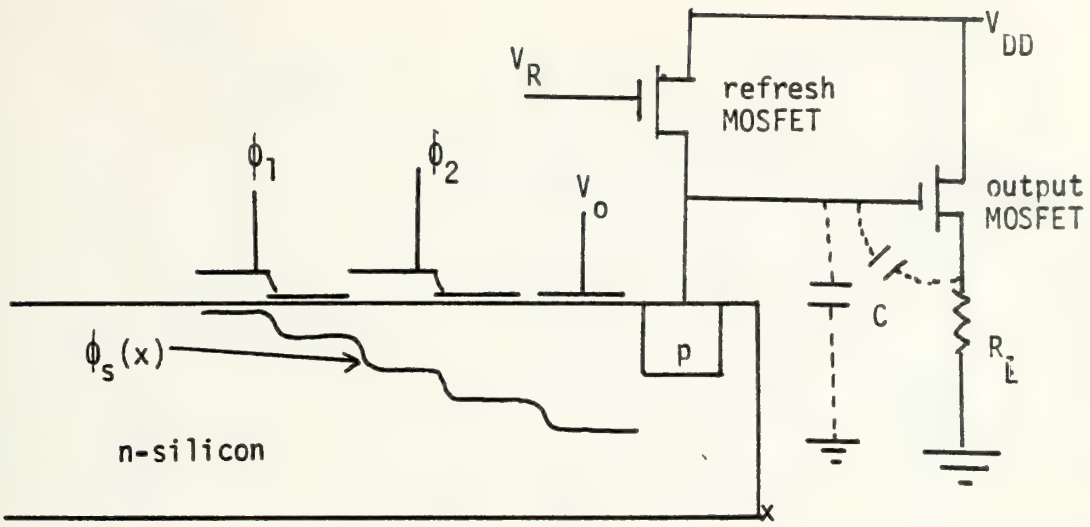


Figure 1.13 Output scheme of the CCD.

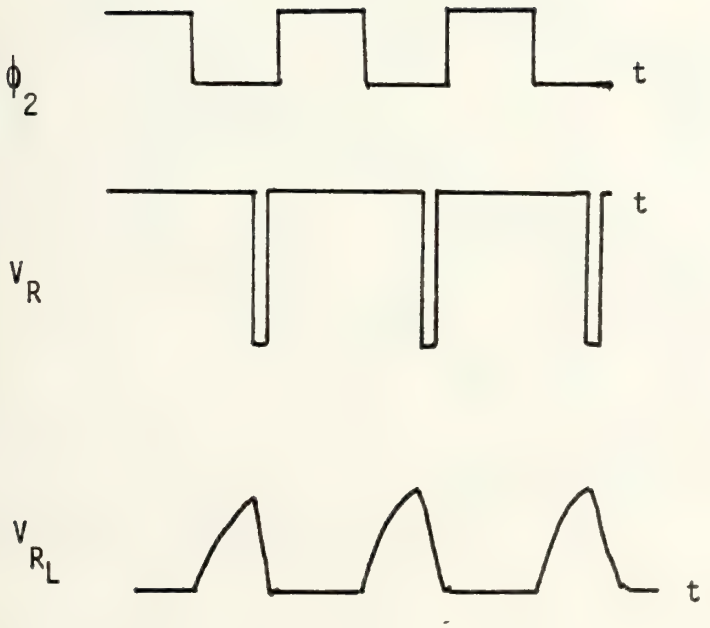


Figure 1.14 Timing diagram of the output stage

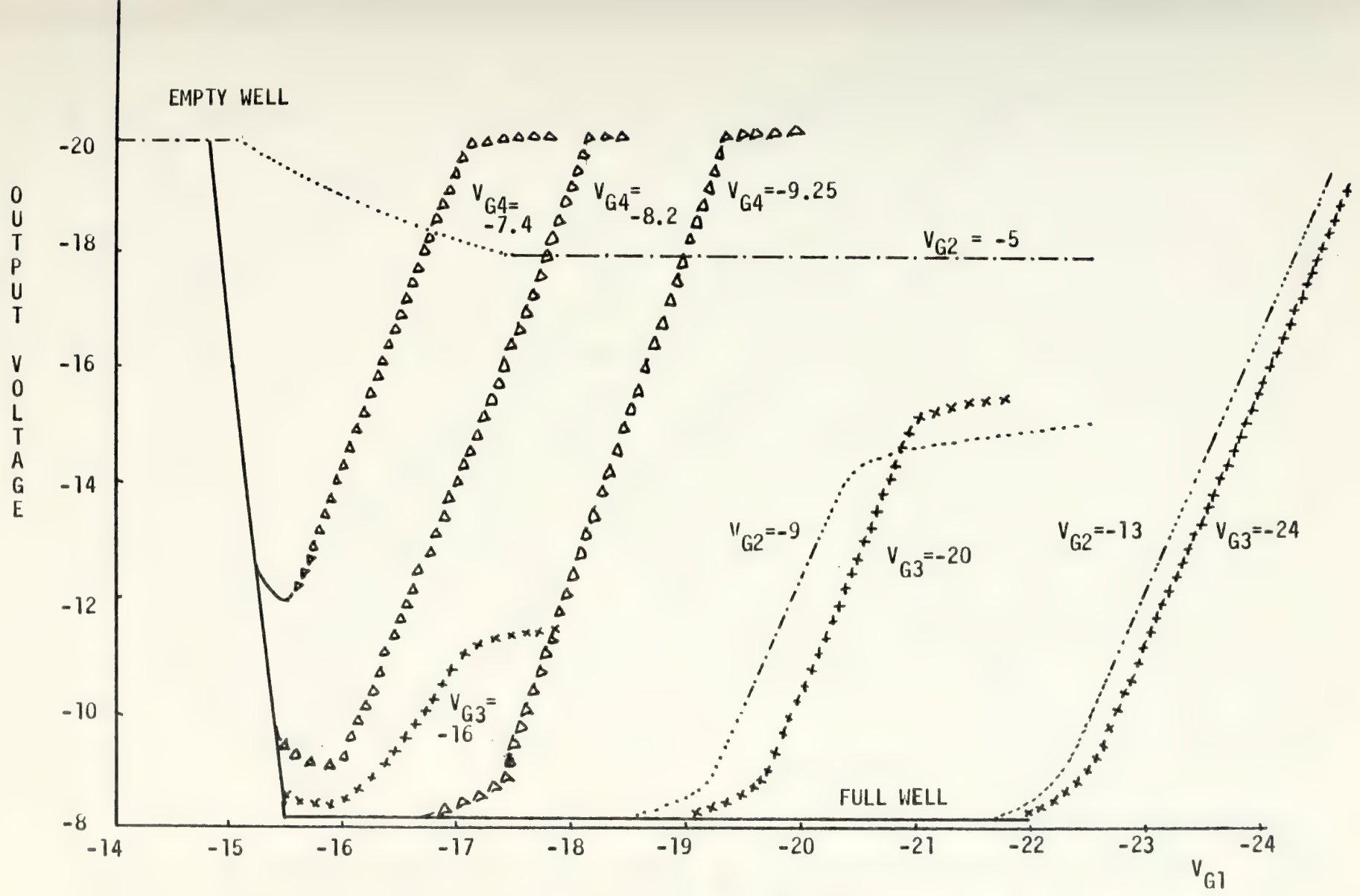


Figure 1.15. CCD operating curves for input through V_{G1}. Dotted curves show V_{G2} as constant parametric value, x curves show V_{G3} as constant parametric value, Δ curves show V_{G4} as constant parametric value.

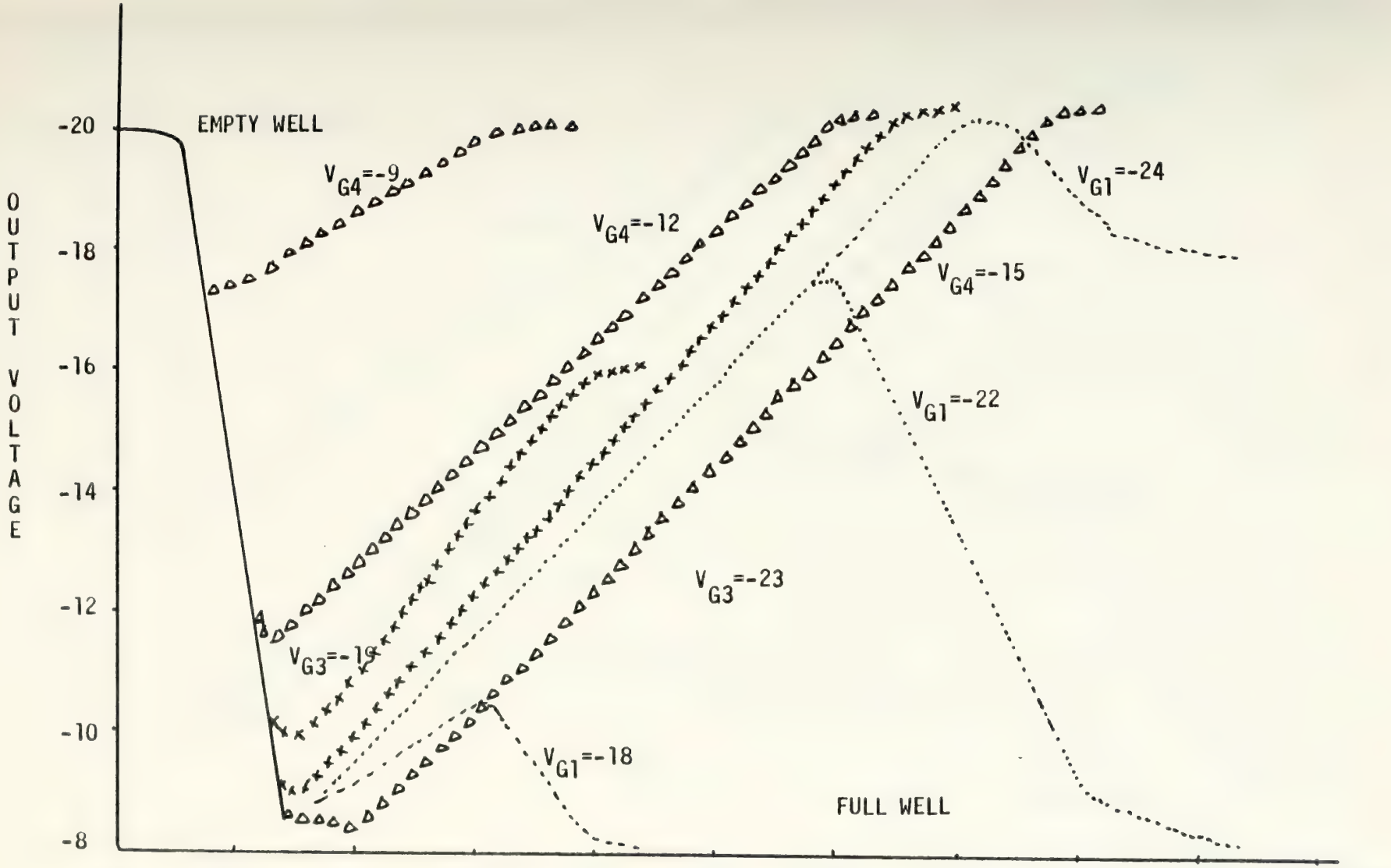


Figure 1.16. CCD operating curves for input through V_{G2}. Dotted curves show V_{G1} as constant parametric value, x curves are for V_{G3}, and △ curves are for V_{G4}

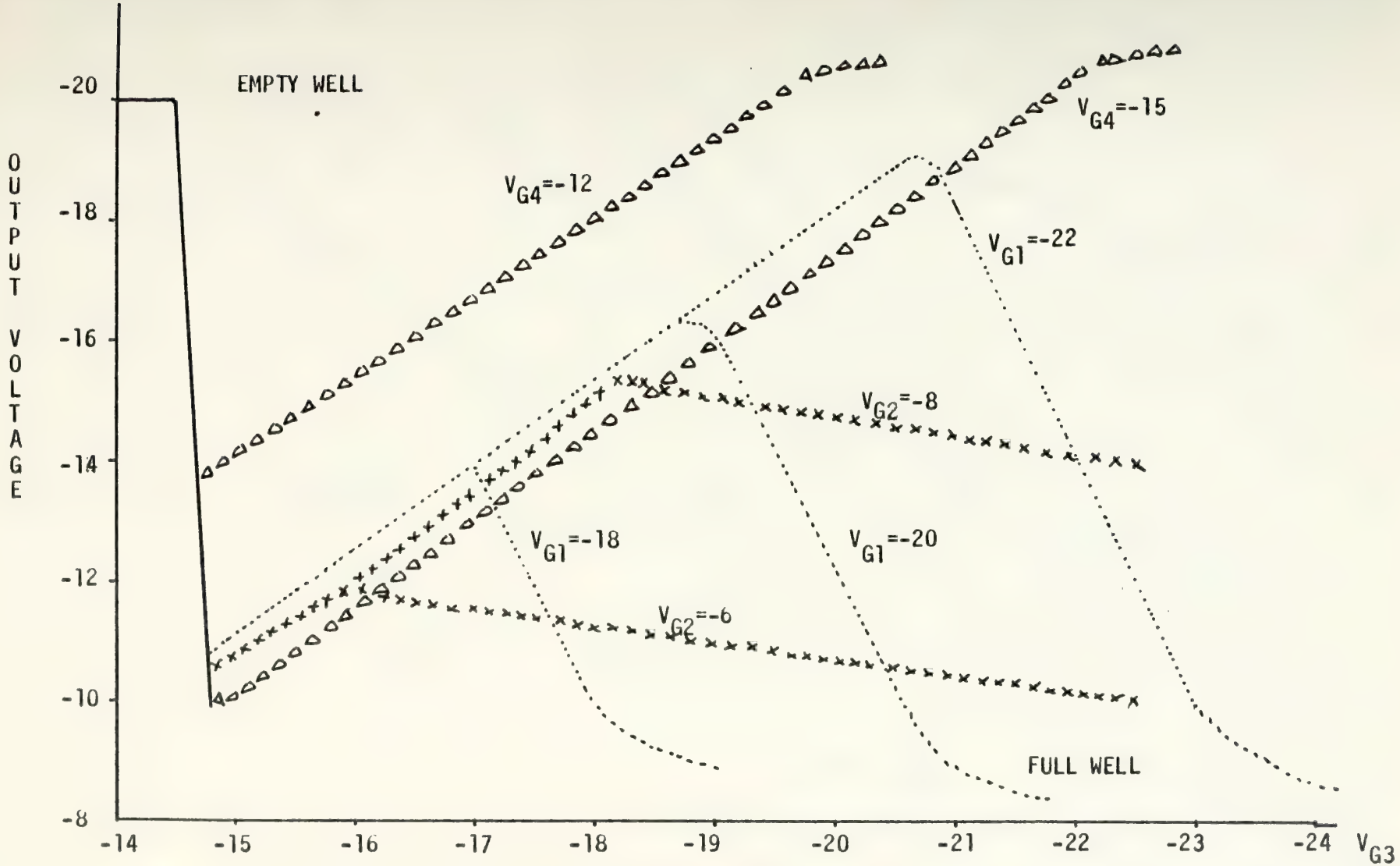


Figure 1.17. CCD operating curves for input through V_{G3}. Dotted curves show V_{G1} as constant parametric value, x curves are for V_{G2}, and △ curves are for V_{G4}.

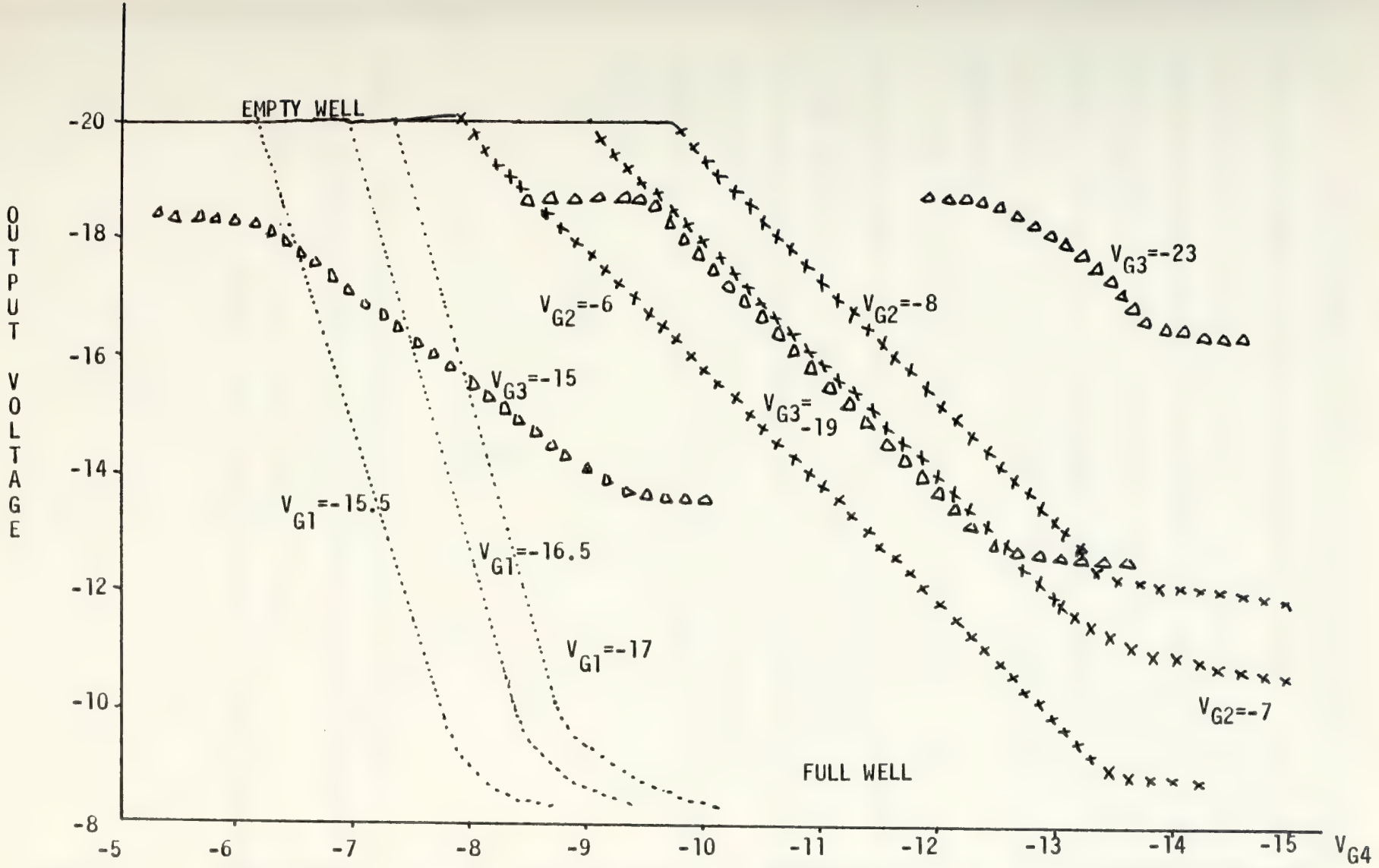


Figure 1.18. CCD operating curves for input through V_{G4}. Dotted curves show V_{G1} as constant parametric value, x curves are for V_{G2}, and $\triangle$ curves are for V_{G3}

the so-called BARRIER mode, which appears to be linear despite theory to the effect that it should vary with the square of the input voltage. The positive sloped part of the characteristic is a noninverting region and is achieved by scuppering. The negative sloped part is an inverting mode, but it, too, is achieved by scuppering. The explanation for the shapes of these curves is taken up in the next section.

One significant aspect of CCD research is the question of what makes the charge move ... that is, does it move from electric fields or by diffusion. As indicated earlier, the first papers assumed that the movement was by diffusion. This has subsequently been shown to be incomplete. As a matter of fact, there are three phenomena which cause movement:

A. Diffusion

$$J_D = - q D \frac{dp}{dx} \qquad (1\text{-}9)$$

where J_D is the current density due to diffusion, D is the diffusion constant for the material, p is the density of holes, and x is the linear distance.

B. Drift under the influence of the Fringing Field:

$$J_{E_f} = q \mu_p p E_f \qquad (1\text{-}10)$$

where J_{E_f} is the current density due to this drift, μ_p is the mobility of holes, q is the charge on one hole, and E_f is the fringing electric field. The fringing field is the field caused by different surface potentials at different regions in the silicon.

C. Drift under the influence of the self-induced field.

$$J_{E_s} = q\mu_p p E_s = -q\mu_p \gamma \frac{dp}{dx} \qquad (1\text{-}11)$$

where J_{E_s} is the current density due to this drift. The total current is the sum of these currents. Gamma, γ, in the last equation is a physical constant depending on the oxide capacitance (and therefore the dielectric constant, the thickness of the oxide, and the doping level), various voltage levels (gate, flat band, threshold), and other constants. It is interesting to note that all three have positive direction at the start of transfer.

The equation for current density is

$$J = q[\mu_p E_f p - (\gamma \mu_p p + D)\frac{dp}{dx}] \qquad (1\text{-}12)$$

Equation (1-13) is the continuity of charge equation.

$$\frac{\partial Q}{\partial t} = - \nabla \cdot J; \qquad \frac{\partial p}{\partial t} = - \frac{1}{q} \frac{\partial J}{\partial x} \qquad (1\text{-}13)$$

By use of this equation and the expression for the current density, we can find a differential equation which relates the time rate of change of hole density to the space rate of change of hole density and the hole density. This is a parabolic differential equation and is not easily solved analytically.

Nevertheless, we can conjecture several facts: First of all, at time zero, the dp/dx term is theoretically infinite, since in all studies, the initial charge density has been assumed to be constant (over x) and zero under the gate to which the transfer is about to be made. We conclude that the self-induced field is very large, relatively speaking. Thus it dominates the current initially. At subsequent time intervals, the fringing field, E_f, becomes the dominant force causing hole movement. As t becomes very large, dp/dx actually changes sign and holes diffuse in the reverse direction, and tend to cancel the drift current caused by E_f. When J = 0 or

$$\mu_p E_f p = (\gamma \mu_p p + D) \frac{dp}{dx} \qquad (1\text{-}14)$$

transfer is complete.

Section 1.3 Experimental Results

In Figures 1.15 through 1.18, we saw the experimental curves relating output to input for our CCD. Given in Figure 1.19 are two curves, one indicating the surface

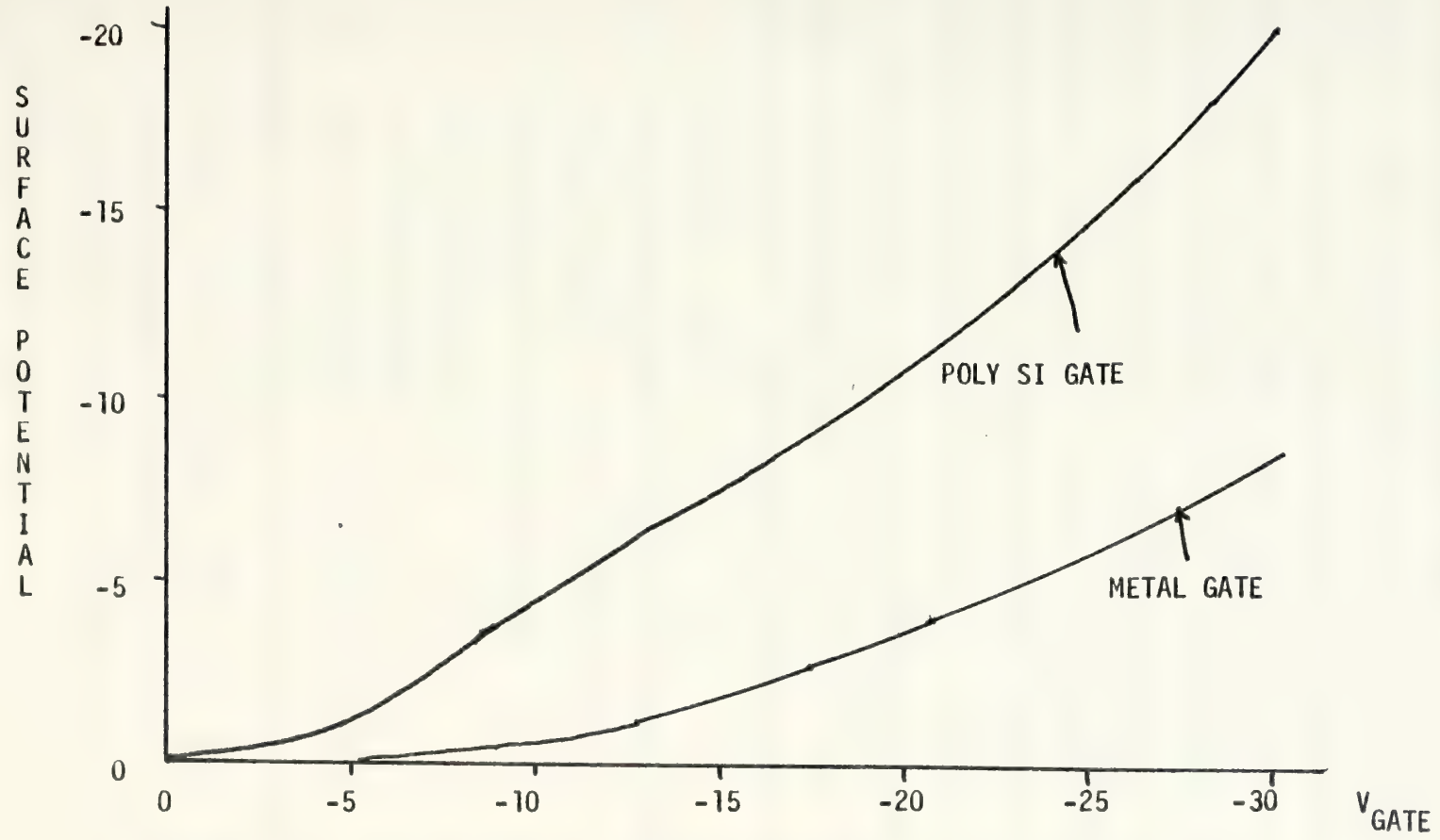


Figure 1.19. Gate voltage vs. Surface Potential for the metal and Poly-Si gates for the TRW CCD utilized in this research.

potential of the silicon substrate under the relatively thick insulator gate (3000 Å), and the other showing the surface potential of the substrate under the thin insulator gate (1000 Å). As expected from equation (1-3), the surface potential as a function of applied gate voltage is much greater for the thin gate than for the thick gate.

It is possible to explain most of the phenomena shown in Figure 1.15 through 1.18 in terms of the shape of the curves in Figure 1.19. Such an analysis is important since it provides an engineering insight into the physical characteristics and limitations of the CCD. The analysis is perfectly general, since it could be made on any CCD as long as the surface potential curves were available.

The CCD consisted of the following physical devices:

1. An input diode
2. Four input gates (of which the second was three times as long as the other three)
3. Eight pairs of transfer gates
4. One output gate
5. One output diode
6. FET's for refresh and output follower

The important parameters of this device are as follows:

1. Oxide thickness under metal gate . . . 3000 Å
2. Oxide thickness under poly Si gate . . 1000 Å
3. Permittivity of Silicon Dioxide . . . 3.9 (relative)
4. Doping level density . . . $1.5 \times 10^{15}/cm^3$
5. Metal gate threshold . . . - 11 v

6. Poly Si gate threshold . . . - 4 v

7. Size of input and transfer gates . 1 mil x 1 mil (except for input gate no. 2 which is 3 mil x 1 mil)

Preliminary Observations (Figures 1.15 through 1.18)

1. The curves show that there are at least three distinct modes of linear operation possible in the CCD, though not all present in any given input scheme. For instance, if the gate voltage of the gate closest to the input diode is varied while another gate voltage is held constant, then there is possible only two operating modes. If, on the other hand, the gate closest to the input diode is kept at a constant voltage and some other gate is used as an input for a variable signal, then three modes are possible.

2. The mode with the greatest gain and least input linear range of operation is the so-called "barrier" mode.

3. The modes with the least gain (but still greater than one) and widest range are the modes which are usually referred to as "scuppering" modes.

4. The "barrier" mode is an inverting mode. Scuppering can produce either inverting or noninverting modes.

5. There is a local maximum (since the scales are shown negative, it is actually a minimum) between the two scuppering linear operating regions.

6. The barrier mode occurs approximately at the same surface potential range for all curves.

Explanation of Experimental Results

Consider Figure 1.15: The gate voltage on G_1 is the abscissa variable while the dotted curves indicate a constant parametric voltage on G_2 with G_3 and G_4 held low. The x curves indicate a constant parametric voltage on G_3 with G_2 and G_4 held low. The Δ curves indicate a constant parametric voltage on G_4 with G_2 and G_3 held low.

Note that the barrier mode appears to occur for the same voltages on G_1 for all input schemes. Since it is extremely important to know the actual surface potentials under the gates as well as the gate voltages, the Figure 1.19 is especially important. If we initially set

$$V_{G1} = -15.0 \text{ v} \qquad V_{G3} = -20 \text{ v}$$

$$V_{G2} = -30 \text{ v} \qquad V_{G4} = -30 \text{ v}$$

and

$$\phi_1 = -10 \text{ and } -24 \text{ v}, \qquad \phi_2 = -10 \text{ and } -24 \text{ v}$$

and the input diode pulse: -12 to -2.5 v
then the Figure 1.15 shows that we are in empty well operation ... that the voltage on G_1 is not low enough to produce the barrier mode much less the scuppering mode. Since we can be sure that the amount of charge transferred to the transfer gate is negligible, the surface potential diagram shows the situation: (Figure 1.20). The actual

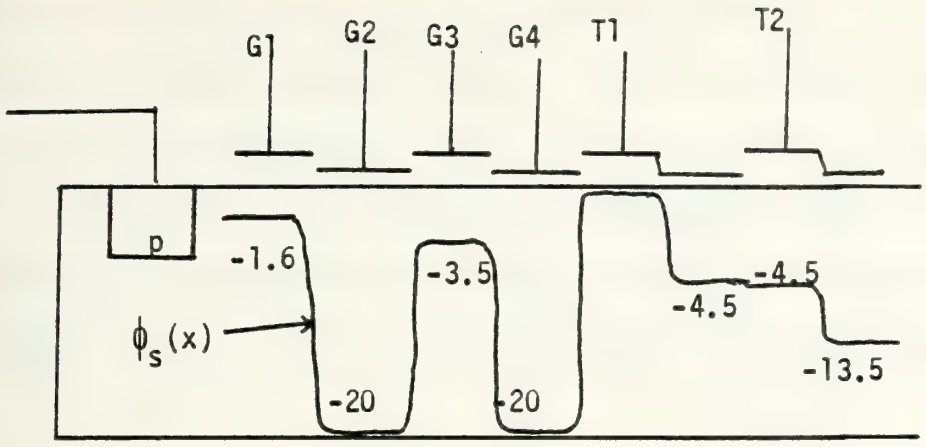


Figure 1.20. Surface potential curves under gates

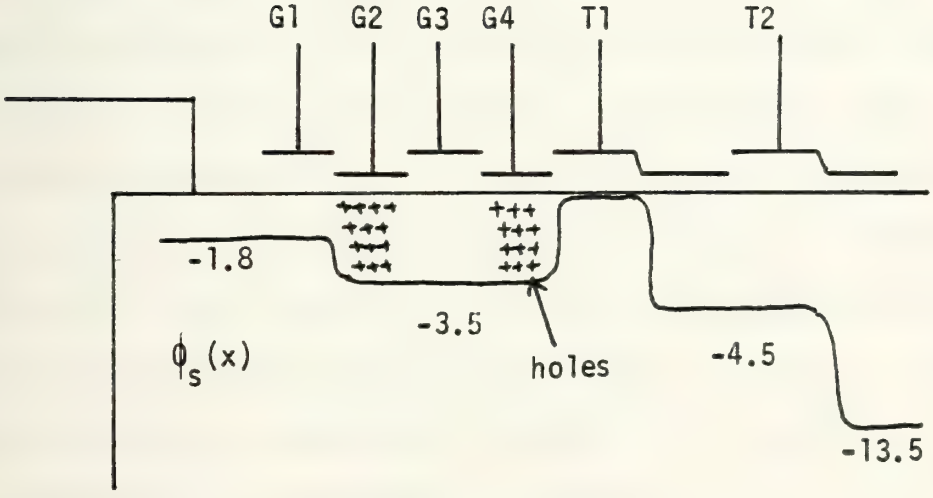


Figure 1.21. Surface potential and stored charge under gates

values for surface potential in Figure 1.20 and succeeding figures have been taken from the curves on Figure 1.19.

The surface potential for the input diode is taken on the n side of the p-n junction. A forward biased p-n junction can be expected to have a voltage drop across it of about 0.7 v. Thus the -1.8 voltage was obtained by

$$- 2.5 - (- 0.7) = - 1.8 \qquad (1\text{-}15)$$

When the input diode pulse does not succeed in forward biasing the diode (as in this case), we do not get a flood of holes into the input gate region. However, the junction IS made less reverse biased, so we do get a few more holes injected (by diffusion) than before. Once a hole has crossed the junction, it can either re-cross into the p-region, or diffuse into the gate regions. We remember that diffusion current is established whenever we have a difference in the concentrations of carriers in two (or more) different regions. Thus, we would expect, for small differences in surface potentials of the input diode and G_1 to obtain some hole injection and subsequent transfer. In analytical terms, the current due to the holes is given by

$$J_p = - q D_p \frac{dp}{dx} + q \mu_p p E \qquad (1\text{-}16)$$

where $\mu_p = 500 \text{ cm}^2/\text{v sec}$ and $D_p = 13 \text{ cm}^2/\text{sec}$.

Thus as long as

$$-\frac{dp}{dx} > \frac{500\ p\ E}{13} \qquad (1\text{-}17)$$

we will get hole injection into the gate region.

When holes reach the region below gate 2, they remain there raising the surface potential. They cannot proceed further because of the great difference (-20 and -3.5 v) in surface potentials between G_2 and G_3, and they cannot return to G_1 for the same reason. They are static charges at this point. However, if more and more holes are injected (as would be the case if V_{G1} were lowered slightly), there will arise the situation of holes "filling" the potential well of G_2. What this actually means is that the holes attracted to the G_2 will finally succeed in raising the surface potential of G_2 to about -3.5 v. At this point, some of the holes will transfer into the G_3 region. Diffusion will bring them to G_4 where they will remain. Of course, the surface potential under G_4 will rise at this point. If there are enough holes injected during the on time of the input pulse, the surface potentials of G_2 and G_4 will both be -3.5 v and will have charge stored. The surface potential of G_3 will also be -3.5 v, but will have negligible charge stored. This condition is shown in Figure 1.21.

If still more holes are injected, they will gather under gates 2,3, and 4 thus raising the surface potential still

further. Finally, at the point of equilibrium, the surface potential under all the gates will be -1.8 v.

When the holes are transferred to the first transfer (or delay) gate, ϕ_1 goes "low" and ϕ_2 goes high. As is shown in Figures 1.20 and 1.21, the empty well surface potentials are

	high	low
For the metal gate:	- 1 v	- 4.5 v
For the poly Si gate:	- 4.5 v	- 13.5 v

where the fact that the high level of the surface potential under the poly Si gate equals the lower level under the metal gate is a coincidence.

The transfer of charge follows precisely the same procedures as the transfer from diode to gate. The holes move, by diffusion and also under the influence of the fringing field and self-induced field, into the transfer gate region. The electric field is strong until the holes reach the area under the poly-Si. But the deposit of charge in this region raises the surface potential. Charge continues to be transferred until the surface potential under the poly Si is about -4.5 v, the same as the surface potential under the metal gate. Thereafter, both regions fill with charge, raising the surface potential even higher. The final value of the surface potential depends, of course, on the amount of charge available for transfer. Thus, if there is a small

amount of charge available, then perhaps no charge at all will be stored under the metal part of the gate. On the other hand, if charge is plentiful, then both regions could have a final surface potential of -1.8 v.

Following the first transfer to T_1, the surface potential diagram might look like Figure 1.22.

In Figure 1.22 we notice that there still is charge stored under G_4 even after transfer. This is because the surface potentials between G_4 and the metal portion of T_1 are now equal. Thus if holes diffuse into T_1, an opposing electric field tends to force them back to G_4. A similar situation is set up under G_2. The charges, permanently stored under G_2 cannot be transferred to G_3 since the transfer would set up an opposing electric field which would tend to drive the charge back to G_2.

Thus far, the mode we have been dealing with has been the barrier mode. This does seem to be a descriptive word to describe the mode, since the junction is only slightly forward biased thus creating a potential barrier to hole injection. That this barrier can be overcome is important to the operation of this mode.

When the gate voltage on G_1 is set lower we immediately get full well operation. With the other voltages set as before, let us assume that $V_{G1} = -17.0$ v. Prior to any charge transfer, the surface potential diagram is shown in Figure 1.23.

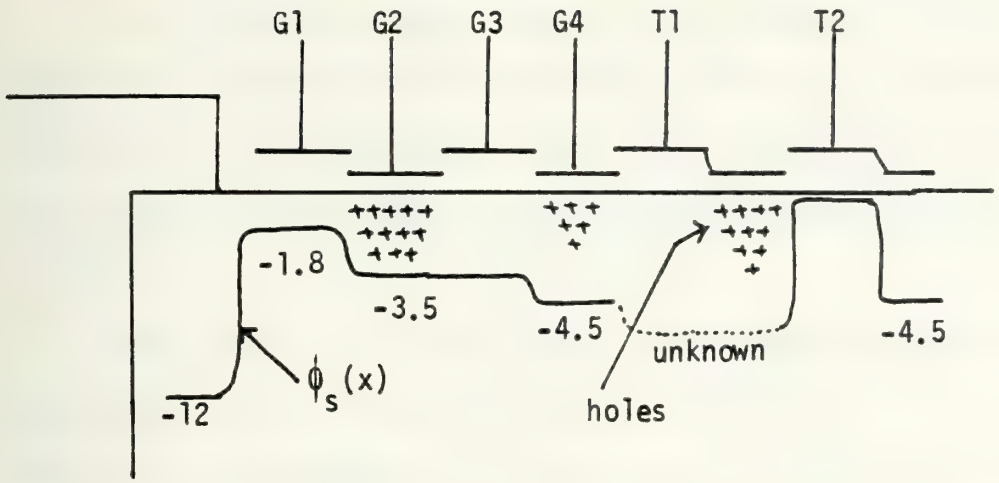


Figure 1.22 Surface potential and charge distribution under gates.

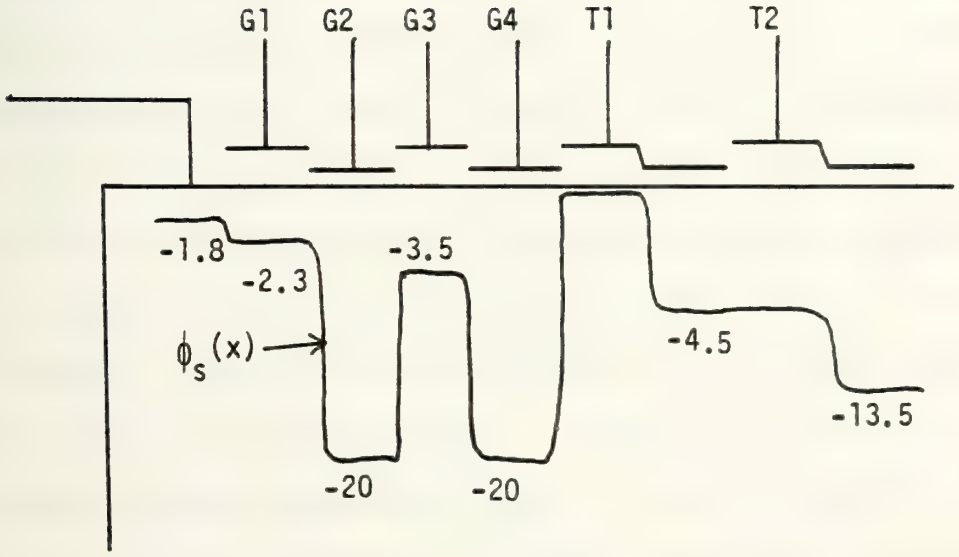


Figure 1.23 Surface potential under gates

Now when the pulse arrives, the junction will be strongly forward biased and holes will be injected easily. They will first accumulate under G_2, then G_4, and finally bring the entire region to a common surface potential of -1.8 v.

When the input pulse falls, the surface potential at the input diode (n-side) will fall to -12.0 v. The charge which is stored under G_1 will drain through the reverse biased diode, leaving G_1 empty, which means that the surface potential under G_1 will return to -2.3 v. Figure 1.24 shows the situation at the instant that the input pulse falls.

In Figure 1.24 it is easy to see that the charge stored under G_1 will return across the p-n junction. It is also apparent that this movement will create an electric field which attracts the holes under gates 2, 3, and 4 to move toward the input diode. When the surface potential under gates 2, 3, and 4 falls to -2.3 v, the transfer of charge across the junction will stop because of the barrier of G_1. This barrier is set up because at some point in the scuppering transfer, further charge movement to G_1 from G_2 will set up an opposing electric field ... that is, G_1 will have a higher surface potential than G_2, thus retarding the hole flow. After scuppering has been completed, the surface potential under all gates will be -2.3 v, and charge will be stored under gates 2, 3, and 4, but NOT under G_1.

The amount of charge transferred to T_1 after ϕ_1 goes low is a fraction of the charge under gates 2, 3, and 4.

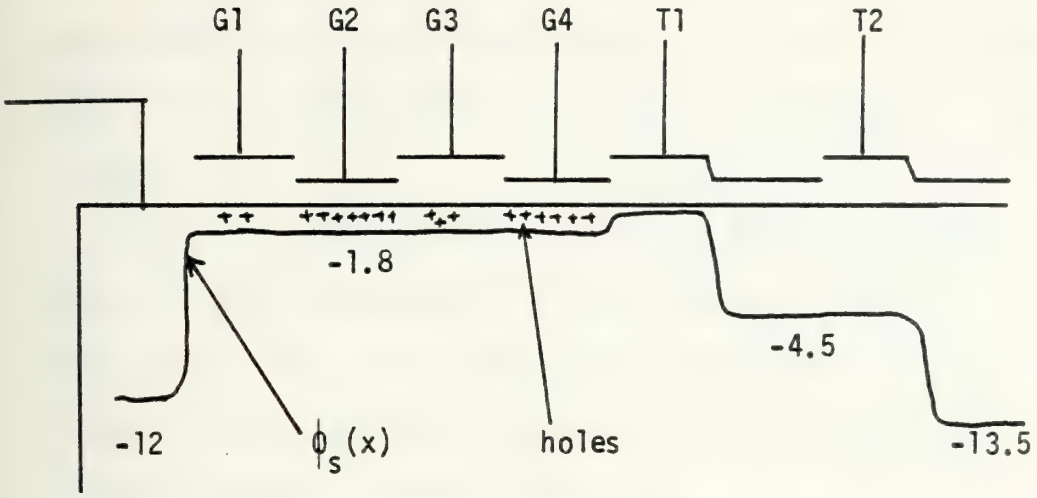


Figure 1.24. Surface potential and charge distribution under gates

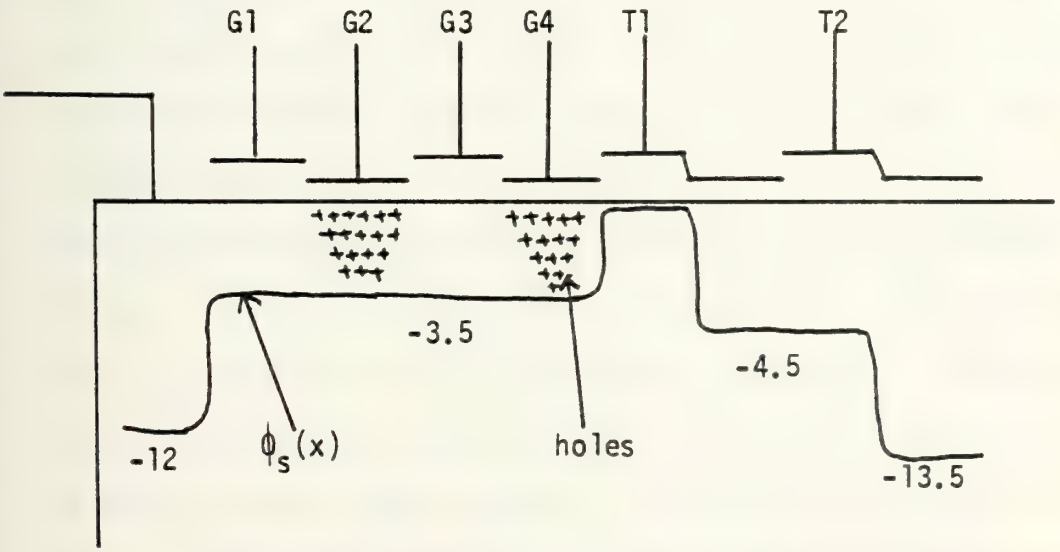


Figure 1.25. Surface potential and charge distribution under gates

If the gate voltage of G_1 is reduced to lower values, LESS charge remains under gates 2, 3, and 4 after the input pulse falls. Thus less charge is transferred to T_1 when ϕ_1 goes low.

Now we set $V_{G1} = -20$ v. As is shown in Figure 1.15, this is near the center of the linear scuppering region. Now assume that the input pulse has arrived, fallen to -12 v, and ϕ_1 is about to go low. We have the surface potential diagram shown in Figure 1.25.

We notice that no charge is stored under G_3. This is because the surface potential of an empty G_3 is precisely -3.5 v. ... the same as it is for an empty G_1. When ϕ_1 goes low, only the portion of charge under G_4 is available for transfer since the surface potential of G_3 represents a barrier to hole transfer from G_2. The figure also shows a previously unexplained phenomenon . . . that is, the amount of charge transferred levels off for decreased values of V_{G1}. This is explained by Figure 1.25. Suppose we let V_{G1} fall to -21 v. The surface potential under an empty G_1, in that case, would be -4 v. The sequence of events is shown by the next series of Figures (1.26, 1.27, and 1.28). In Figure 1.28 we see the final arrangement of charge prior to the falling of ϕ_1 causing transfer. The only charge available for transfer is that under G_4. Further, we see that an additional decrease in V_{G1} will have no effect at all on the amount of charge stored under G_4. This amount of charge MUST be enough to cause the surface potential

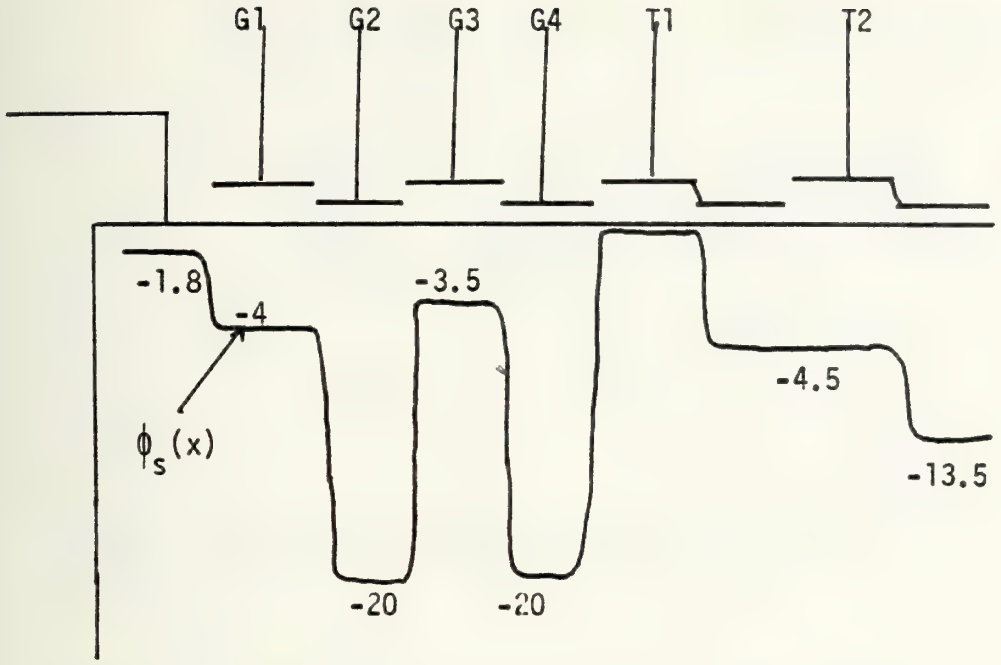


Figure 1.26. Empty well surface potential under gates

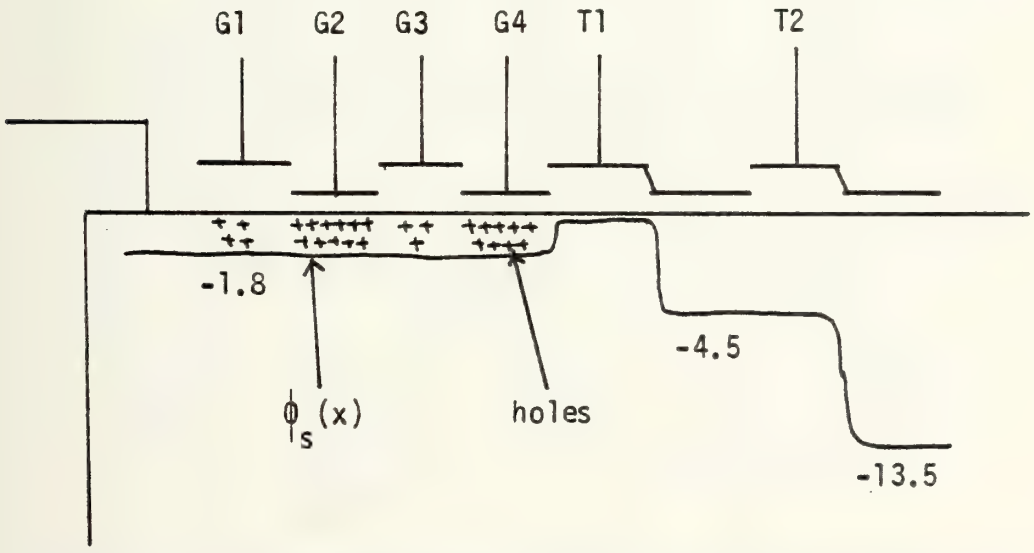


Figure 1.27. Surface potential after hole injection but before the input pulse falls.

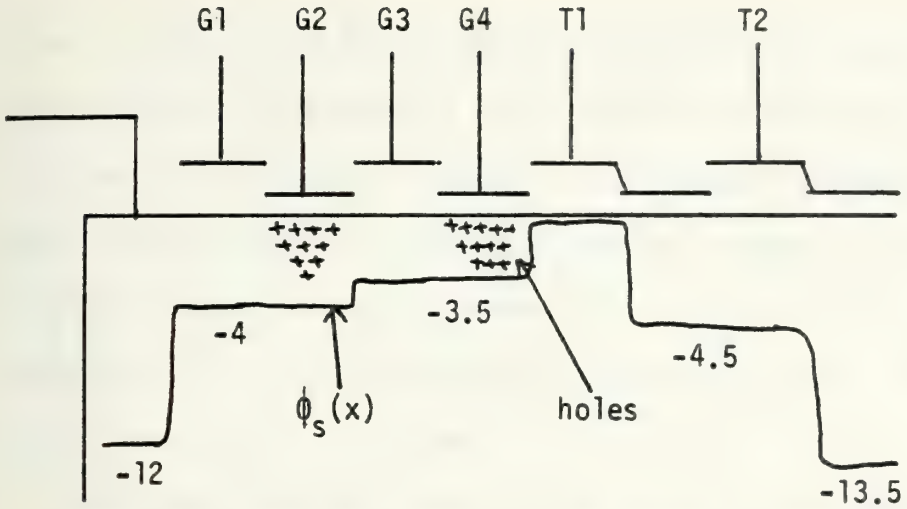


Figure 1.28 Surface potentials and stored charge after input pulse falls and excess charge is drained.

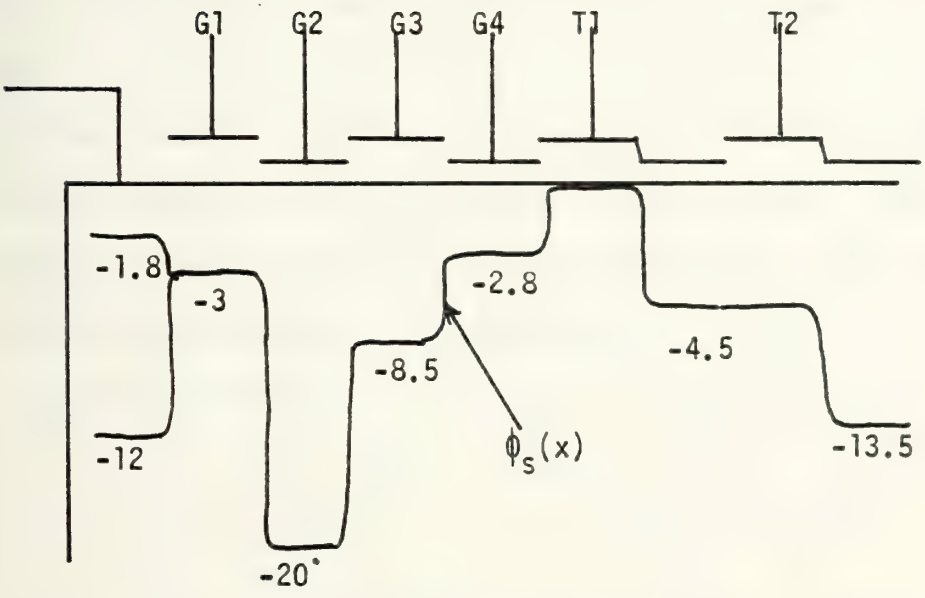


Figure 1.29 Surface potentials for empty wells. Note that the surface potential under G4 is higher than that under G1. This fact will be important.

under G_4 to be -3.5 v. More charge, of course, cannot remain under G_4 since it will raise the surface potential and cause an electric field to influence its movement leftward. The fact that this is the only charge available for transfer, does not, of course, mean that we have full well operation. On the contrary, we see that the amount stored is not enough for full well operation.

Moreover, we see that no matter how low the gate voltage of G_1 goes, we can never recover empty well operation with the current settings of V_{G3}, V_{G2}, and V_{G4}. Thus, we get a leveling off region . . . which is confirmed by the curves.

A similar phenomenon takes place for variations in V_{G1} and a constant parametric voltage, V_{G2} (gates 3 and 4 held low).

The situation of holding V_{G4} as a constant parametric voltage and varying V_{G1} is also interesting. The figure shows that we first get the barrier mode as explained above, and then immediately get scuppering.

Suppose we set

$$V_{G1} = -19 \text{ v} \qquad V_{G4} = -7.5 \text{ v}$$

and keep gates 2 and 3 low. Figure 1.29 shows the empty well surface potentials. Notice that the surface potential under G_4 is higher than that under G_1. This fact will be important.

Following the usual chain of events, the input pulse causes hole injection which raises the surface potential of all gates to about -1.8 v. When the input pulse falls, the excess charge is drained off leaving the surface potentials of gates 1, 2, and 3 at -3.0 v, and the surface potential of gate 4 at -2.8 v. In addition, it is to be noted that there is charge stored ONLY under gates 2 and 3, as shown in Figure 1.30.

Now when ϕ_1 falls, there is NO charge available for transfer since no charge is stored under G_4, and there is a barrier between the charge stored under G_3 and the first transfer gate. Thus we get empty well operation. This is consistent with the curves in Figure 1.15.

Now if V_{G1} is raised to -16.5 v, the empty well surface potential of G_1 is slightly higher than that of G_4. Thus after scuppering, we have the following charge and surface potential diagram in Figure 1.31.

A small amount of charge is available for transfer according to Figure 1.31. When the charge is transferred, the resulting surface potential of gates 2, 3, and 4 can be NO LOWER than -2.8 v. Thus, only the charge differential between the surface potentials of -2 and -2.8 v can be transferred to T_1.

As V_{G1} is increased, more charge is transferred.

As V_{G1} is decreased, less charge is transferred. These conclusions are consistent with the curves on the Figure 1.15.

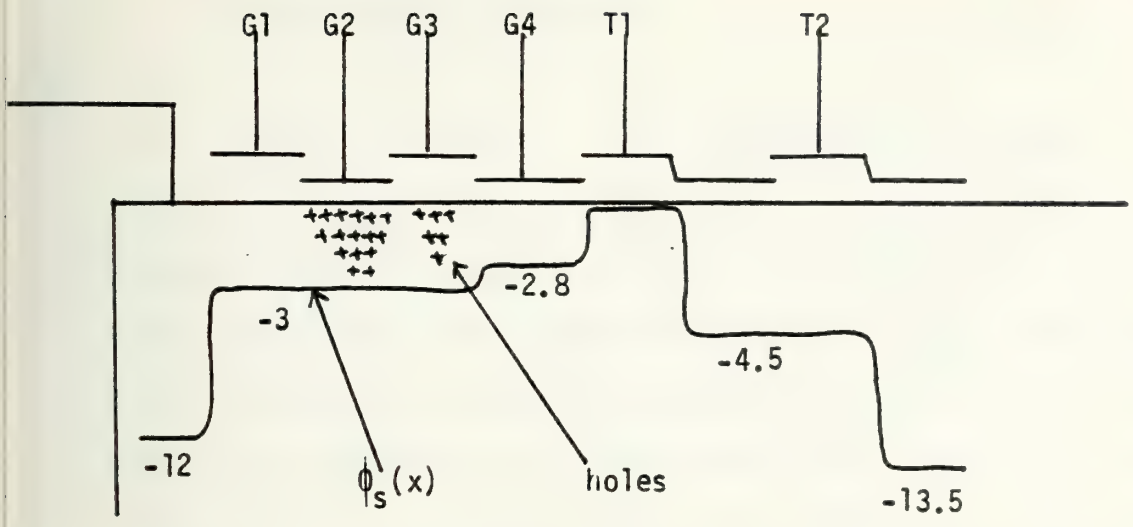


Figure 1.30. Surface potential and charge distribution after diode pulse falls.

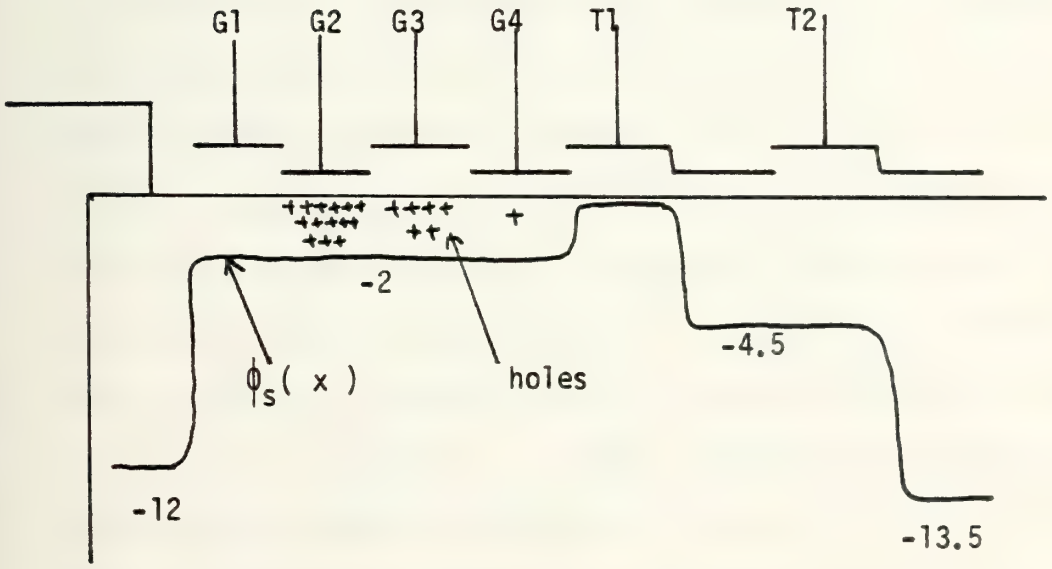


Figure 1.31. Surface potential and charge distribution under gates

If we make the settings

$$V_{G2} = -5v \qquad V_{G3} \text{ and } V_{G4} \text{ held low}$$

and vary V_{G1}, the curves tell us that there is a gentle slope from empty well operation to almost full well operation. We find the surface potentials for this case from the Figure 1.19 curves. We notice that the surface potential under G_2 is about -1 v. Thus there is a weak electric field to inhibit hole movement. The diffusion component of current, however, is strong since $\frac{dp}{dx}$ is large. Thus we expect to find a small amount of charge transferred and this is, of course, what we get.

There is one more phenomenon, perhaps not so obvious from the curves, which should be considered. It was stated that the same phenomena occurred when V_{G2} was held constant as when V_{G3} was held constant. That is, there is a leveling off effect which cannot be changed by variations in V_{G1}. There is a certain (depending on the constant value selected for V_{G2} and V_{G3}) amount of charge below which it is impossible to go by varying V_{G1} alone. However, the amount of charge can be reduced (to zero, it turns out) by reducing the constant parametric value of V_{G2} or V_{G3}.

Suppose we set

$$V_{G1} = -30 \text{ v} \qquad V_{G2} = -13 \text{ v}$$

with gates 3 and 4 held low.

The surface potential (empty well) diagram is shown in Figure 1.32.

In the usual manner, the input pulse injects holes into the gate region, raising the surface potential to about -1.8 v. Then the pulse falls to -12.0 v and the excess charge in the gate region drains across the reverse biased p-n junction. We are left with the following charge and surface potential diagram (Figure 1.33).

It would seem, at first glance, that the charge stored under gates 3 and 4 would be available for transfer. Such is not the case. Notice the surface potential under the metal part of the transfer gate which has ϕ_2 impressed on it. This surface potential is -4.5 v. This level is the same level as attained by ϕ_1 when it goes low. Thus when ϕ_1 does fall, we have the following surface potential and charge diagram (Figure 1.34).

No charge can be transferred from G_4 to T_1 because there is already a potential barrier preventing hole injection. Thus we get empty well operation as predicted and confirmed by the curves. Notice that the condition described is a function of the surface potential under the metal gate when ϕ_1 goes low. This, in turn, is, of course, a function of the low level of ϕ_1, which for our purposes was kept at -24 v. For the above case, if this level were reduced, the value of V_{G2} necessary for empty well operation would also be reduced. In other words, we would get a wider input

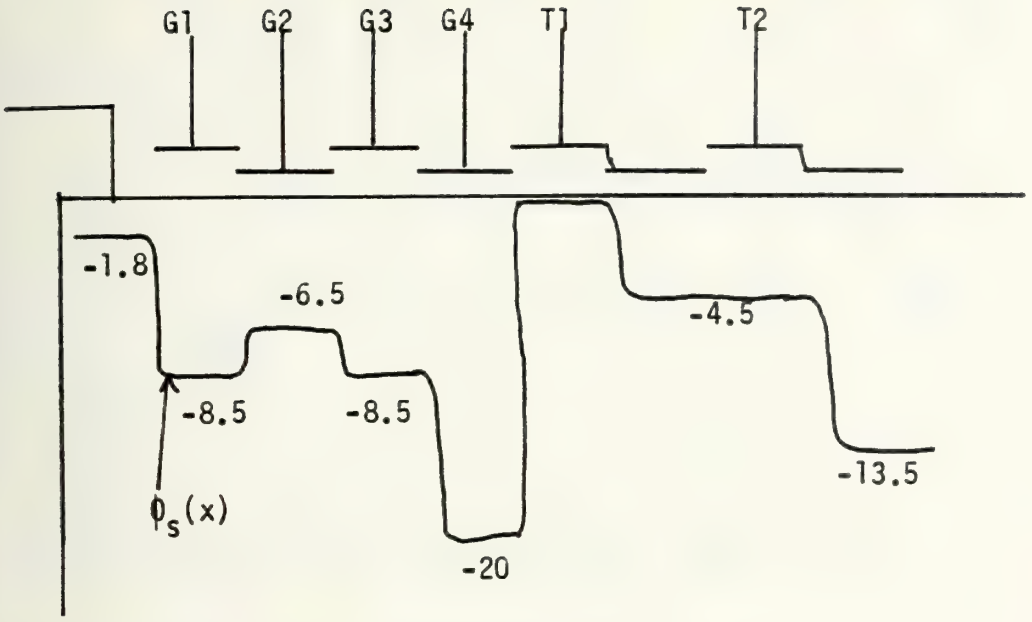


Figure 1.32. Surface potential for empty wells.

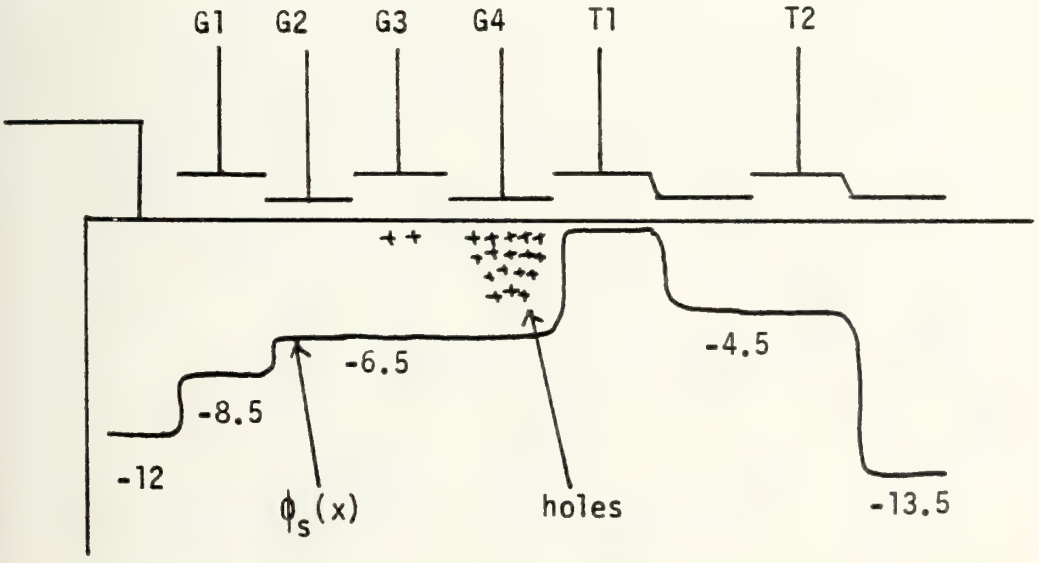


Figure 1.33. Surface potential and charge distribution after input diode pulse falls and excess charge drains across p-n junction.

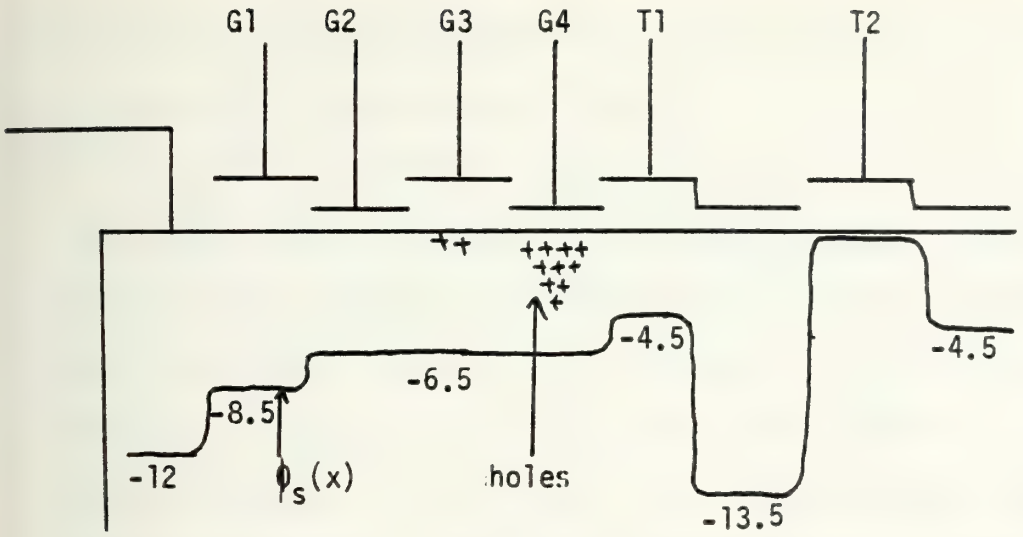


Figure 1.34. Surface potential and charge distribution after ϕ_1 falls and ϕ_2 rises. Notice that no charge transfer is made.

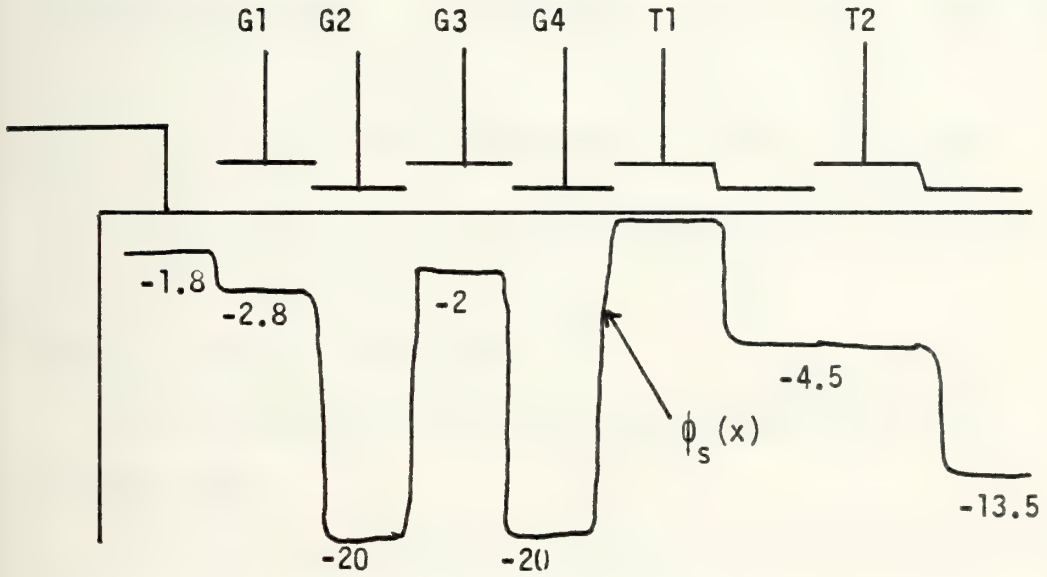


Figure 1.35. Empty well surface potential

range and lower gain for this input scheme if the low level of ϕ_1 could be less than -24 v.

We now turn our attention to Figure 1.17 which shows the variation of V_{G3} with constant $V_{G1,2,4}$. Immediately we notice some differences between this figure and Figure 1.15. Most obviously, we see in two cases, three separate linear regions. The first is, of course, the aforementioned barrier mode. But the latter two seem to have about the same gain, except one inverts and the other does not. There is also a local maximum (voltage) between the two latter modes.

Let us bypass the barrier mode as having been explained adequately above. Let us make the following settings:

$$V_{G1} = -18 \text{ v (constant parametric value)}$$

$$V_{G3} = -16 \text{ v (variable)}$$

and put gates 2 and 4 low.

The initial surface potential diagram is shown in Figure 1.35.

In the usual manner, the input pulse injects holes into the gate region thus raising the surface potential in the gate region to -1.8 v. When the input pulse falls to -12.0 v, the excess charge drains through the p-n junction leaving the surface potential and charge distribution as shown in Figure 1.36.

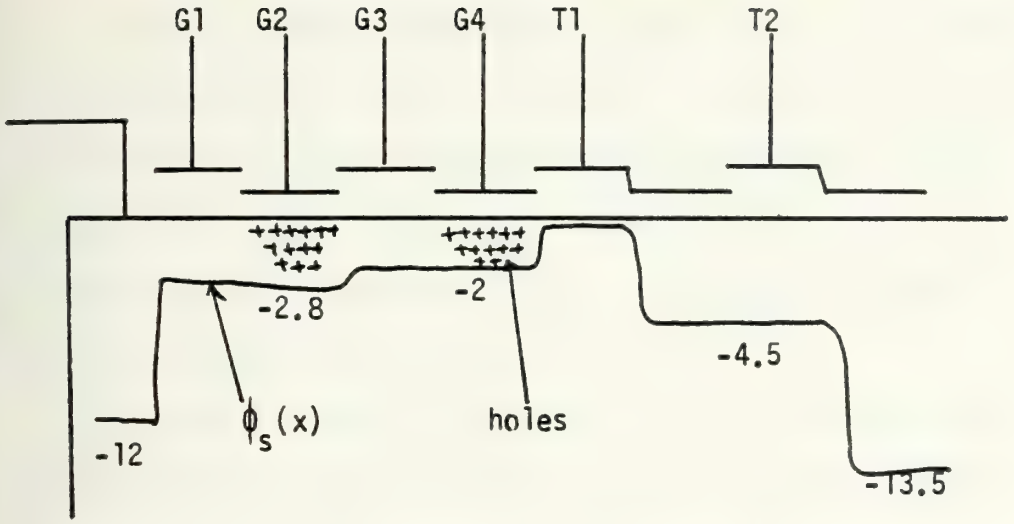


Figure 1.36. Surface potential and charge distribution under gates

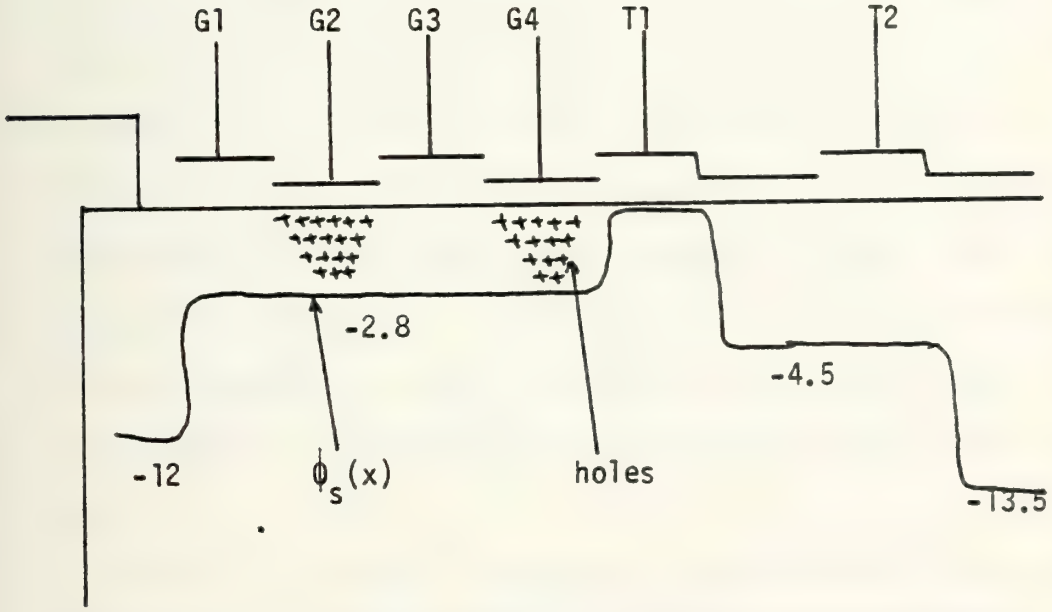


Figure 1.37. Surface potential and charge distribution at local maximum

When ϕ_1 falls, the charge under G_4 is available for transfer to T_1. If V_{G3} were to be reduced, it would tend to reduce the charge transferred since more would be drained across the input junction. Thus we have a noninverting mode which corresponds to that seen on the figure (dotted curve labeled $V_{G1} = -18$ v).

However, there is a limit to the reduction of V_{G3} which can be done and still decrease the charge transfer. If we reduce V_{G3} to -18 v, the empty well surface potential of G_1 is the same as for G_3. After the falling of the input pulse, this new situation is shown in Figure 1.37.

This represents the local maximum of charge transfer. If V_{G3} is further reduced, say to -19 v, then after scupper-ing, there will be THE SAME AMOUNT of charge stored under G_2 and G_4 as is stored in the local maximum condition, but there will be an ADDITIONAL charge stored under G_3. And this charge is available for transfer. Thus, for further reductions in V_{G3}, we get MORE, not less charge transfer. This corresponds to the curves in Figure 1.17 which shows this maximum and inverting/noninverting modes on either side.

Figure 1.38 shows the charge distribution and surface potential of the CCD where V_{G3} has been decreased beyond the local maximum.

Comparisons were made between the actual numbers obtained on the curves and the corresponding surface potentials as predicted by TRW. Correspondences were quite close.

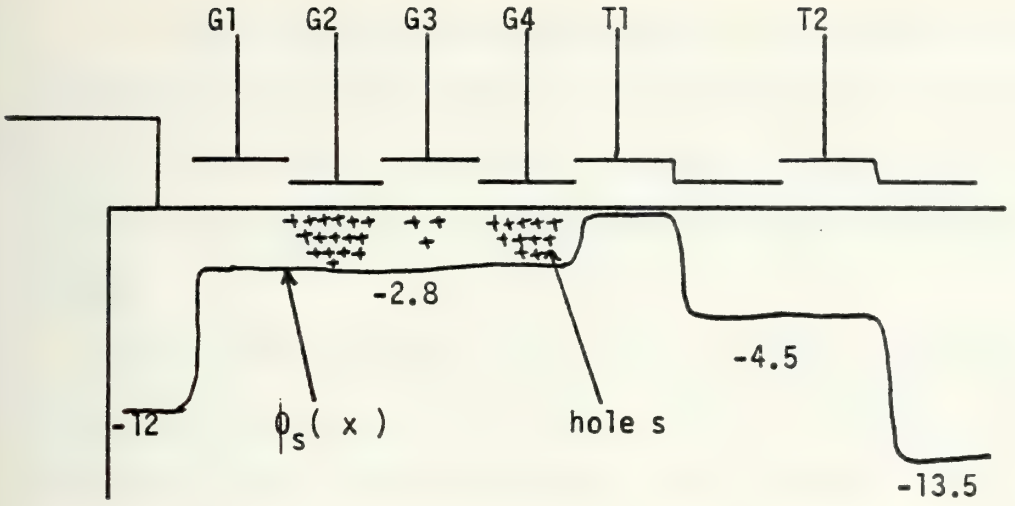


Figure 1.38. Surface potential and charge distribution where V_{G3} has been decreased beyond the local maximum

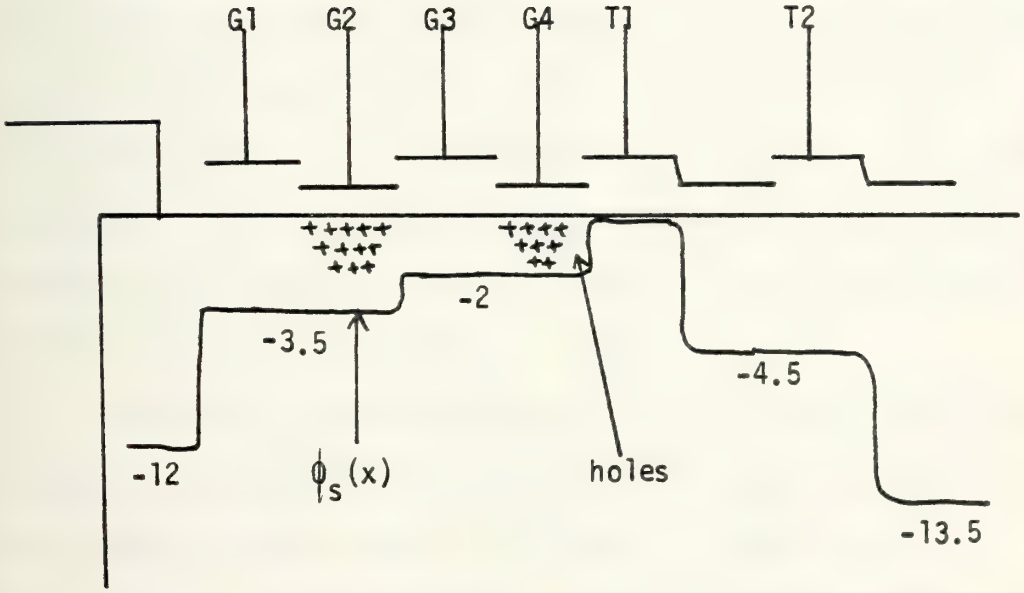


Figure 1.39. Surface potential and charge distribution under input gates.

Of course, the same explanation can be made for the appearance of three modes and a local maximum for varying V_{G3} and constant parametric values of V_{G2}. The numbers are a bit different, but, otherwise, the explanation is the same.

One conclusion that we reach is that the various constant values of, say V_{G1}, should not have any effect on the output so long as the maximum has not been reached (by V_{G3}). That is, the curves, V_{G1} = -18 v vs. V_{G3}; V_{G1} = -20v vs. V_{G3}; and V_{G1} = -22 v vs. V_{G3} should all be the same until V_{G3} is reduced to the extent of reaching the local maximum. This is confirmed by the curves in Figure 1.17.

The reason this should be so can be seen by considering the charge available for transfer . . . it is only the charge stored under G_4. And this charge is not a function of V_{G1} (as long as V_{G1} is low enough to insure that the surface potential of G_1 is below the surface potential of G_3). Figure 1.39 shows a typical charge distribution and surface potential diagram for this situation.

Suppose V_{G1} were decreased. It is clear that this would affect the amount of charge stored under G_2, but would have no effect on the amount of charge stored under G_4. Thus, the curves should be common, and it is seen, they are.

Some of the other parameters which we wanted to consider are the various values of gain, linear range, and dynamic range. Table 1.1 shows the values of gain as a function of the gate used as the input gate and the gate used as the

		constant parameter gate voltages			
		G1	G2	G3	G4
INPUT	G1	-	5.70	5.79	4.95
	G2	2.02 (-4.4)	-	2.09	2.05
	G3	1.3 (-4.14)	1.3 (-.269)	-	1.33
	G4	-7.2	-2.0	-1.73	-

Table 1.1. GAIN (scuppering modes) as a function of input gate and constant parameter gates

constant parametric voltage gate (both in scuppering modes). Inspection of the table shows that for noninverting scuppering, gain changes little as long as the same input gate is used. The reasons for this can be found in the same type of analysis as was undertaken to explain the shape of the operating curves. Consider the noninverting scuppering mode of operation with gate 1 serving as the input gate. At the time of transfer into the first transfer gate, the charge available for transfer lies spread throughout the input region: in the potential wells of gates 2, 3, and 4 (but not gate 1). The amount of charge in this region can, of course, be increased by reducing one of the gate voltages (2, 3, or 4). However this modification changes the gain very little because very little of the increased

charge finds its way into the first transfer gate. As Figure 1.40 shows, the charge in the region under gates 2, 3, and 4 will increase with a reduction in one of the voltages, but the charge available to the transfer gate is only the charge which is excess to a surface potential of -4.5 v. By excess to a surface potential of x volts, we mean that portion of the charge, which, if it were absent, would lower the surface potential to x volts. Thus, the available charge changes very little, if at all.

On the other hand, if the input were to be impressed onto gate 2, then the charge available for transfer would lie under gates 3 and 4 in the noninverting scuppering mode. Figure 1.41 shows a typical charge profile for this situation. Since there is less charge available for transfer, we might expect that the gain would be smaller, since more voltage would be required to draw more charge into the region. But, once this situation exists, it is apparent that the actual transfer does not depend on HOW the charge was introduced. Thus the gain should be the same for all three cases of constant parametric voltage. And table 1.1 shows that this is correct.

When we come to input by gate 3, we find that the only charge available for transfer lies in the region under gate 4. Since there is still less charge to be transferred than in the first two cases, we expect that the gain would continue to decrease, and this is borne out by experimental fact. Figure 1.42 shows a typical charge profile for this case.

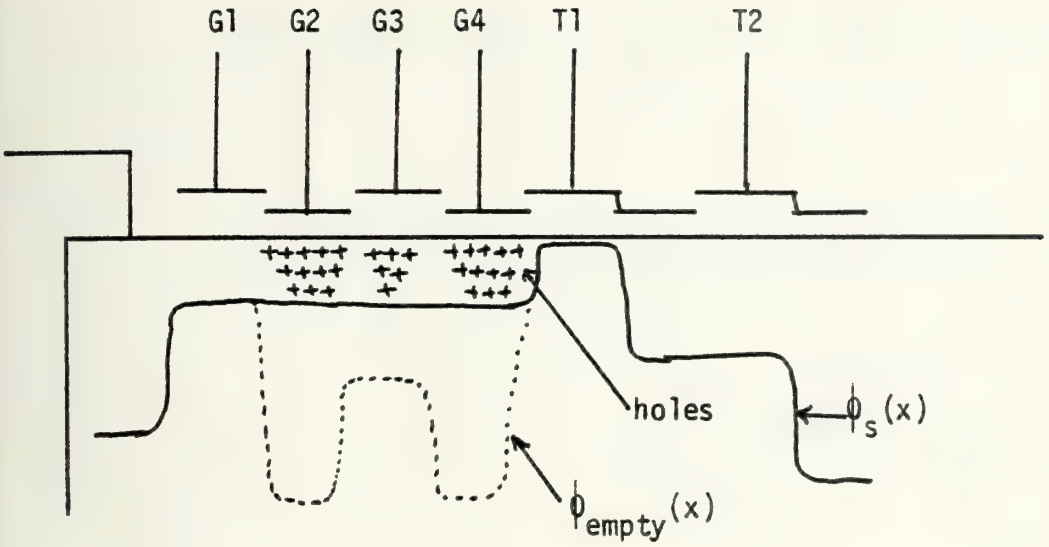


Figure 1.40. Surface potential and charge distribution for a typical noninverting scuppering mode with the input impressed upon gate 1. Charge available for transfer lies under gates 2, 3, and 4. The dotted line is the empty well surface potential

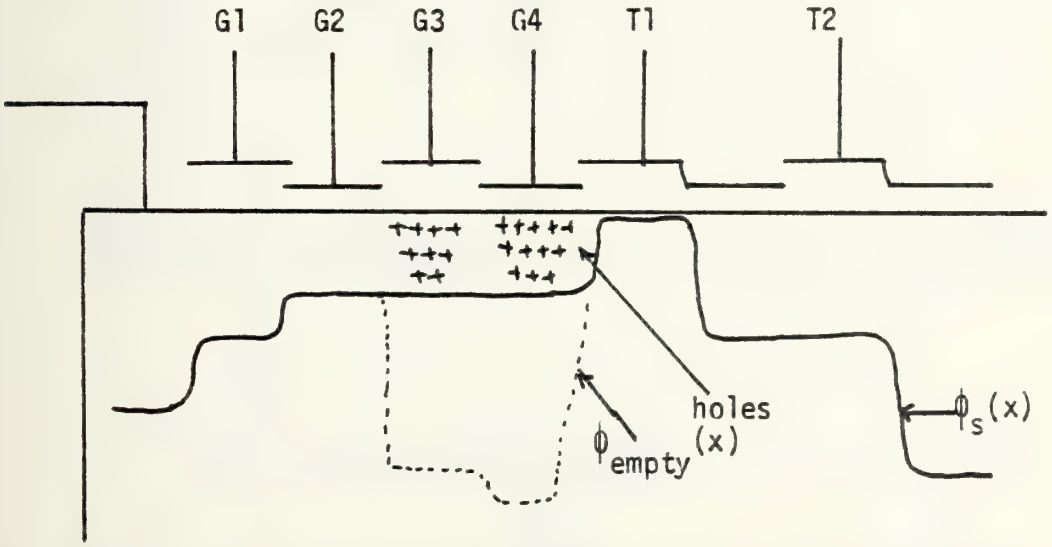


Figure 1.41. Surface potential and charge distribution for a typical noninverting scuppering mode with the input impressed upon gate 2. Charge available for transfer lies under gates 3 and 4. The dotted line is the empty well surface potential.

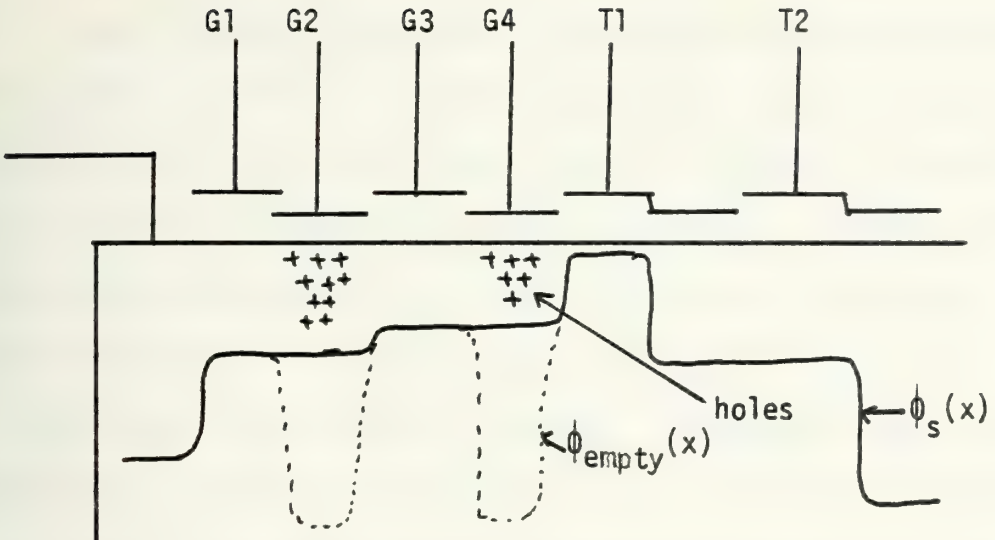


Figure 1.42. Surface potential and charge distribution for a typical noninverting scuppering mode with input on gate 3.

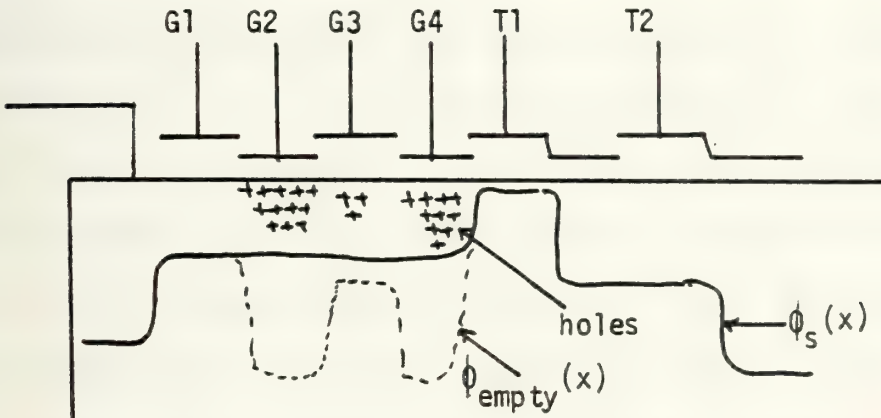


Figure 1.43. Surface potential and charge distribution for the inverting scuppering mode with input on gate 4, V_{G1}=parameter.

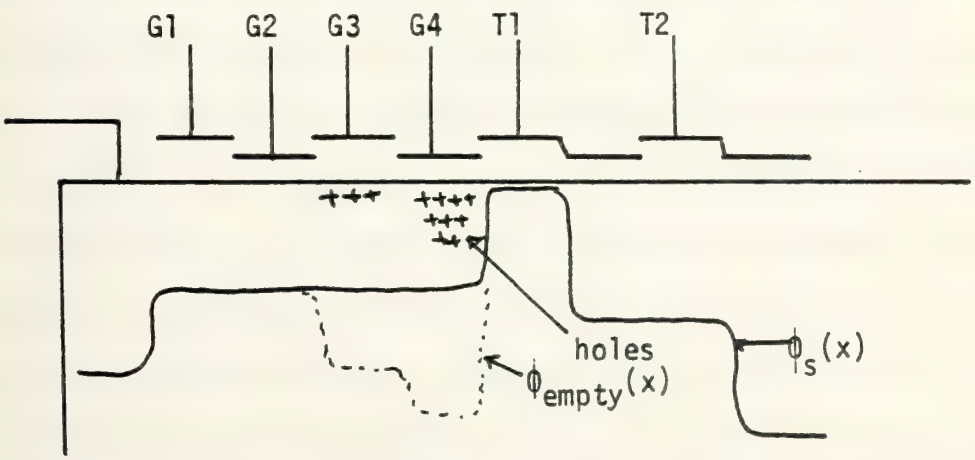


Figure 1.44 Surface potential and charge distribution for the inverting scuppering mode with input on gate 4, V_{G2} = parameter

There is no noninverting scuppering mode for the gate 4 input. A little reflection can make this clear. There is no way (since these are linearly sequential gates in space) to introduce MORE charge into the transfer gate as the gate voltage on gate 4 is INCREASED. Increasing the gate voltage on 4 reduces the amount of charge under gate 4 and, more specifically, reduces the amount of charge excess to a surface potential of -4.5 volts. Thus reducing the gate 4 voltage is the only way to obtain more charge transfer. This is a case of inverting gain. As for the magnitude of the gain, one can see that it is definitely NOT constant. . . that it is a strong function of the gate used for the constant parametric voltage. The reasons for this can be seen by a look at the charge profile situation existing for each case. When V_{G1} is used as a parameter, as shown in Figure 1.43, the charge available for transfer lies under gates 2, 3, and 4. All the charge in this region which is excess to the surface potential under gate 4 is available for transfer. Moreover, as Figure 1.44 shows, if V_{G2} is used as the constant parameter, the charge available for transfer is limited to the region under gates 3 and 4. Similarly, if V_{G3} is used as a parametric voltage, then the charge available for transfer lies only under gate 4. Thus, one would expect that the magnitude of the gains would be decreasing as the parameter voltages moves from gate 1 to 3. And this is the result we find in table 1.1.

The other inverting scuppering modes follow similar rules except for the strange situation of impressing the input onto gate 3 while using V_{G2} as the parametric voltage. The magnitude of the gain is quite small (-0.269) and no satisfactory reason has been found for it. One conjecture is that the marginal charge available for transfer by reducing V_{G3} is so small as to be almost negligible. But it is not clear why this should be so. The charge available for transfer can be shown to lie under gates 2, 3, and 4 in this situation, not unlike others we have investigated. However, in this case, the CCD behaves abnormally. Such a small gain is not very useful for design engineers, and thus it probably has no practical significance, yet an explanation for its "gain mutation" would be enlightening.

Other parameters which are important are linear and dynamic ranges. The input linear range can be seen from Figures 1.15 through 1.18. The widest linear range is about 7.5 volts on Figure 1.17 using V_{G3} as the input and V_{G4} as the constant parameter voltage. This is surely adequate for most applications.

The dynamic range for this CCD was approximately 60 db.

We found that the shape of the output pulse (full well operation) of the CCD became considerably poorer as frequency was increased. The limit for proper operation was around 75 kHz, although it could be operated at higher frequencies with an attendant loss in gain. We consistently used a conservative 20 kHz clock frequency.

The maximum delay time can, of course, be found from a calculation of storage time. Empirically, we found that we obtained full well operation in all modes at frequencies lower than about 12 kHz. This is a result of the dark current filling the transfer gates during the "low" time of the clock pulses. This corresponds to a storage time of 83 msec and a total maximum delay time of about half a second. Naturally, this calculation is approximate since dark current causes considerable distortion even if it does not totally fill the well. Better CCD's have been made whose delay time is measured in seconds. The watchwords of these CCD's are low inefficiency, long registers, and a small dark current.

Section 1.4 Conclusions

It is important to note that in order to undertake this analysis, many voltage levels had to be defined including the levels on the refresh FET, clock voltages, the input diode bias and pulse levels, plus the output resistance (load) value, the width of the input pulses, and the width and timing of the refresh pulses. Such a large number of parameters is, of necessity, a disadvantage when attempting to utilize the device for systems applications. Manufacturers of CCD's must simplify the operation of the CCD such that several of the voltages can be tied together at some mutually appropriate level. It should be possible by design of the geometry and the thicknesses of the oxide levels to

insure proper operation of the device with, say $V_o = V_{DD}$. The user ought to be able to connect the unused input gates to V_{DD} as well.

Some simplicity in timing circuits also ought to be designed into the chip. The clock pulse train driving each transfer gate is a square wave and the optimum positions (in time) for arrival of the input diode and refresh pulse are determined, completely, by the transfer clock square wave. It is possible in today's high semiconductor technology to expect CCD makers to provide one pin for the clock and have internal circuitry to provide inversion (for the opposite clock phase) and to provide timing signals for the input and refresh pulses.

A serial-in, serial-out CCD designed for scuppering, would be left with the following pinouts:

1. V_{DD}	3. G_1	5. G_3	7. output
2. Ground	4. G_2	6. G_4	8. clock (master)

If proper design of the geometry was achieved, and if the characteristics of the device could be predicted (as we think they can), then this dream might prove to be a reality. As it is at present, there are so many voltages and timing problems that it is only with difficulty that the devices can be made to operate at all.

This difficulty was especially troublesome when it came to designing and correctly biasing a second order filter with

two CCD's serving as the sampled analog delay elements. Two CCD's had to be found which were sufficiently "close" in their operating characteristics that certain bias voltages could be used for each device. This was achieved, but due to the drift of the levels, a close watch on the output waveform of the CCD had to be maintained at all times, since a drift in the wrong direction of any number of bias and clock levels caused the filter to malfunction.

This last experience brings us to the other urgent need: the need for reproduceability. Industry consistently tries to sell products whose parameters they can maintain within a set of tight specifications. Clearly, the CCD's we used in this research (some experimental models from TRW SYSTEMS) did not meet this criterion. Nevertheless, as will be shown in the next chapter, we were able to design both first and second order filters and test their performance.

CHAPTER TWO

UTILIZATION OF THE CCD AS DELAY ELEMENTS IN RECURSIVE FILTERS

2.0 Introduction

The contents of this section serve to extend the study from the CCD itself to how the sampled-analog delay element may be used as a circuit element. The following subjects will be addressed.

Section 2.1. The recursive first order comb filter

Section 2.2. Experimental results and theoretical calculations of representative first order comb filters.

Section 2.3. The recursive second order comb filter

Section 2.4. Design methods for a second order sampled-analog comb filter.

Section 2.5. Experimental results and theoretical calculations of representative second order comb filters.

Section 2.6. Conclusions.

SECTION 2.1. THE SAMPLED ANALOG RECURSIVE FIRST ORDER COMB FILTER

2.1.1. The basic characteristics of a sampled-analog recursive first order comb filter are straightforward. The filter has, at most, one zero and one pole. The transfer function is most conveniently written in the z-domain:

$$H(z) = \frac{a_o + a_1 z^{-1}}{b_o + b_1 z^{-1}}$$

where

$$z^n = e^{jn\omega T}$$

which has the result that for real frequencies of operation, the amplitudes of the transfer function can be found by inspecting the functional values on the unit circle of the z-plane. By its recursive nature, feedback is present. Feedback has a significant effect on the time response, a characteristic which plays a major role in applications. Recursive filters have the widely assumed bad property of having a unit pulse response which is infinite in duration. Naturally, transversal filters (those without feedback) have no such disadvantage.

Since there is only one pole and one zero, the job remains, then, to choose the "best" positions on the z-plane for the singularities. With only one pole and one zero,

such a task is not especially difficult since the coefficients used to set the poles and zeros are numerically related to the position of the singularities in a simple manner. That is, the design engineer in the laboratory can change the potentiometer performing the b_1 multiplication and know, rather intuitively, what is happening to the pole on the z-plane.

The first method of first order recursive filter design consists of simply transforming a known transfer function in the s-domain to an "equivalent" one in the z-domain. Several types of continuous filters are widely known and documented. It is a simple matter to choose a first order Butterworth filter, for instance, and apply a standard z or bilinear z transformation to its s-plane transfer function. The resulting function in z is the transfer function of the sampled-analog filter.

As an example, consider the low pass transfer function in the continuous time domain:

$$H(s) = \frac{A}{s + s_i} \qquad (2\text{-}1)$$

If we apply the standard z transform to this transfer function, the usual procedure is to find the inverse La Place transform (h(t)), and sample it at intervals of, say, T seconds. This produces

$$h'(n) = h(nT)$$

The Z transform of h'(n) is

$$Z(h'(n)) = Z(h(nT)) = \sum_{n=0} h(nT) z^{-n} \qquad (2\text{-}2a)$$

$$= \frac{A}{1 - e^{-s_i T} z^{-1}} \qquad (2\text{-}2b)$$

The form shown in Equation (2-2b) makes it easy to pick off the proper values of the coefficients for the sampled-analog recursive filter:

$$a_o = 1 \qquad b_o = 1$$

$$a_1 = 0 \qquad b_1 = - e^{-s_i T}$$

where A is the overall gain of the filter.

The important property of the standard z transform is, of course, that the transformation from s-plane to z-plane maintains an invariant impulse response. By this is meant that the sampled impulse response of the continuous time filter is equal to the unit pulse response of the sampled-analog filter. Such an invariance is implied by Equation (2-2a). This property is important wherever time domain response is a critical specification.

More often, however, the frequency response transfer function is the dominant specification and this is where the standard z transform is deficient. The sampling process of the standard z transform inevitably produces a finite amount of aliasing into frequency response of the sampled-analog

filter. The reason for this can be seen by observing that the frequency content of the continuous filter is infinite ... that is, the amplitude of the frequency transfer function is nonzero for arbitrarily high frequencies (the only zero is at infinity). When such a signal is sampled at some finite sampling frequency, f_s, the Nyquist sampling theorem predicts that the segment of the transfer function greater than $f_s/2$ will cause "aliasing" or "foldover". This aliasing produces amplitude and phase distortion in the output and prevents precise recovery of the original signal no matter what signal processing techniques are utilized. This same aliasing in the sampled-analog filter causes the amplitude of the transfer function to deviate from that of the s-plane transfer function at all points along the $j\omega$ axis. Thus, the frequency response of the sampled-analog filter is different from the continuous time filter. This result is inevitable and exists in all filters derived by the standard z transform. Of course, the deviation may be small and the application may well tolerate such deviation. But the deviation is always present.

When this deviation becomes large, it is possible for important parameters to shift outside the region of acceptability. For instance, the cutoff frequency[1] of a sampled-analog

---

[1]Defined as the frequency for which the response of the filter is down 3 db from the response at dc (for low pass filters).

filter designed with the standard z transform will always be higher than the cutoff frequency of the analog filter from which it was derived. In extreme cases (extreme, but not unusual) the cutoff frequency may cease to exist at all. One example of this is shown in Figure 2-1. The frequency response of the sampled-analog filter never takes on a value of 0.707H(0). Thus, there is no cutoff frequency. Clearly, such a filter, designed for low pass use, is worthless.

The rule of thumb by which sampled-analog filters are designed by the standard z transform method is that the sampling frequency must be quite large with respect to the major frequency content of the continuous-time transfer function. To illustrate, we consider Figures 2.2a and 2.2b. In Figure 2.2a, the sampling frequency is clearly large enough to prevent any significant aliasing from occurring in the sampled analog filter since it is much higher than the range of frequencies where the transfer function is large. However, in Figure 2.2b, the sampling frequency lies in the range of frequencies where the transfer function is significant. Thus, we would expect substantial distortion of both amplitude and phase because of excessive aliasing.

Another method for finding an "equivalent" sampled analog filter for a given continuous time filter is to use the bilinear z transform. Here the transformation into the z-domain is done directly from the s-domain with no intermediate step of finding the inverse LaPlace transform of the

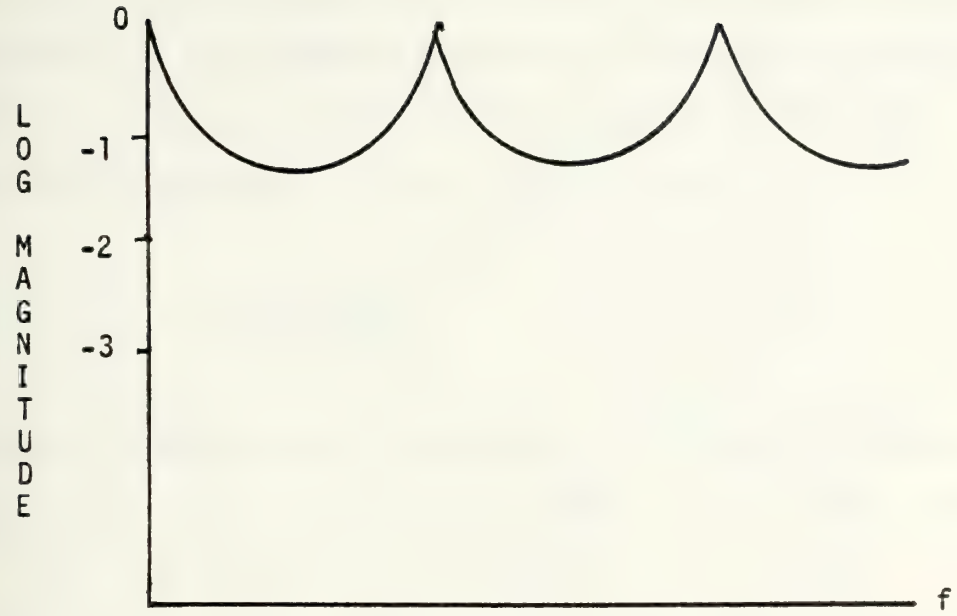


Figure 2.1. Example of a sampled analog filter whose transfer function is such that there is no cutoff frequency...no frequency for which $|H(f)| = .707\,|H(0)|$.

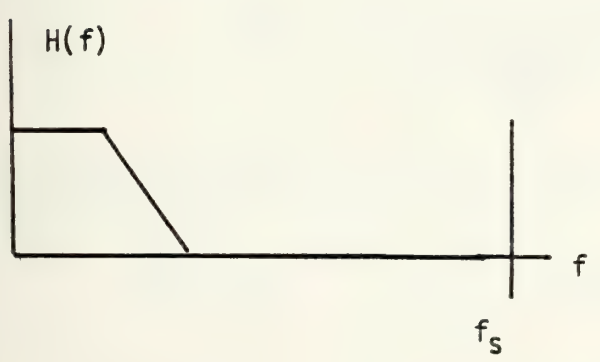


Figure 2.2a

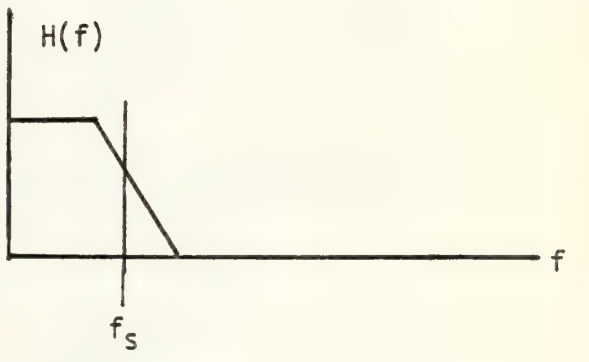


Figure 2.2b

Figure 2.2. Examples of the selection of sampling frequency. When a large (compared to the frequency transfer function domain) sampling frequency is chosen, very little aliasing takes place. However, if too low a sampling frequency is chosen, aliasing can make the filter useless.

analog filter transfer function. Any transformation which preserves the requirement that real frequencies be transformed onto the unit circle of the z-plane will work. However, it is customary to use the simplest:

$$s = \frac{2}{T}\frac{z - 1}{z + 1} \tag{2-3}$$

Such a substitution requires only algebra to derive the transfer function such that the coefficients corresponding to a_o, a_1, b_o, and b_1 can be found.

For our low pass example given for the standard z transform method, we simply make the substitution defined by Equation (2-3) and clear of fractions to obtain

$$H(z) = G\frac{1 + z^{-1}}{1 + \frac{s_iT - 2}{s_iT + 2}z^{-1}} \tag{2-4}$$

This form permits us to find the values of the coefficients:

$$a_o = 1 \qquad b_o = 1$$

$$a_1 = 1 \qquad b_1 = \frac{s_iT - 2}{s_iT + 2}$$

The single most important property of the bilinear z transform is that the resulting sampled analog filter assumes all the functional values that the analog filter does. That is, the range of the functions are precisely equal. Now, this does not mean that the equivalent values of the transfer

functions occur for the same frequencies. As a matter of fact, in general, this is not the case. This change in range vs. domain is called "frequency warping" in the literature, but it is simply because we are dealing with two different functions. Design rules for such filters dictate that a desired cutoff frequency in the sampled analog domain must be found by solving the following equation:

$$f_c = \frac{1}{2\pi T} \cos^{-1}\{\frac{-2b_1}{1 + b_1^2}\} \qquad (2\text{-}5)$$

Conversely, if our cutoff frequency is known (a design specification), then we can calculate b_1 by the following relation:

$$b_1 = \frac{-1 \pm \sin 2\pi f_c T}{\cos 2\pi f_c T} \qquad (2\text{-}6)$$

As a result of the property which forces the ranges of the sampled analog and continuous filters to be equal, we must have at least one zero in the frequency range $0 \leq f \leq f_s$, since, as always, the transfer function is periodic. This zero can be seen to occur at $f_s/2$, since this is the point for which $z = e^{j\omega T} = -1$. Thus the shape of the transfer function is distorted from the shape of the analog filter in order to require the zero to occur. This is, of course, a good example of what causes the "frequency warping" phenomenon.

In general, the relation between the continuous frequency variable and the sampled analog frequency variable for which equal amplitudes of the transfer functions occur is given by

$$\omega_c = \tan \frac{\omega_d T}{2} \qquad (2\text{-}7)$$

where ω_c is the radial frequency in continuous time and ω_d is the radial frequency in the z-domain.

Nothing has been said about the unit pulse time response of the sampled-analog filter. As a matter of fact, the unit pulse time response of the sampled-analog filter does not, as in the standard z transform method, equal the sampled impulse response of the continuous time filter. For those applications where invariance of the time response is a necessity, clearly, the bilinear z transform cannot be used.

The importance, however, of the equivalence in ranges of the analog and sampled-analog transfer functions cannot be overemphasized. It is this property which affords the designer the luxury of obtaining an appropriate analog filter (Butterworth, etc.), applying the bilinear z transform, and being sure that the amplitude values of the analog filter will be faithfully reproduced in the sampled analog filter. There is never a question of the existence of a cutoff frequency. If it exists in the analog filter, it must exist in the sampled analog filter (albeit, at a different frequency). If, as some applications require, there is a range of frequencies for which the amplitude

response is down, say, 30 db from the maximum response in the analog filter, there must exist a range of frequencies in the sampled analog domain which exhibits the same behavior. This fact has enabled us to emphasize design by bilinear z transform in much of our experimental work.

For reference purposes, Table 2-1 has been included. The table shows the relation between the low pass and high pass analog filters and their relatives in the sampled analog domain. Both bilinear and standard z transforms are performed on these filters with the resulting coefficients as shown. We will explore these effects in greater detail in Section 2.2.

TABLE 2-1

DESIGN METHOD	FILTER TYPE	a_o	a_1	b_1
Standard z	Low Pass	1	0	$-e^{-\omega_c T}$
	High Pass	1	0	$e^{-\omega_c T}$
Bilinear z	Low Pass	1	1	$\pm \frac{\omega_c T - 2}{\omega_c T + 2}$
	High Pass	1	-1	$\pm \frac{\omega_c T - 2}{\omega_c T + 2}$

2.1.2 Comb Filters

The basic characteristics and the design rules of first order filters are so simple that it is apparent that anytime a filter of higher order can be put into the first order form, it is a wise thing to do. For instance, if we have a sampled-analog filter of, say, eighth order, and if it can be shown to be equivalent to some first order filter, then, we need only to consider the first order filter, and, at the conclusion of our design, make the proper translation into eighth order.

In our research, we have used a serial-in, serial-out charge coupled device as our analog delay element. The specific chip we have used contained eight delay stages so that eight samples of the input were stored in the CCD at all times. The filter which we designed utilizing this analog delay line consisted of multiplicative coefficients at the input and output of the CCD, followed, of course, by high input impedance summers. The filter set-up is shown in Figure 2.3. Notice that the coefficients, a_8 and b_8, multiply the output voltage by an appropriate value prior to summation. The transfer function of this filter is

$$H(z) = \frac{a_o + a_8 z^{-8}}{1 + b_8 z^{-8}} \qquad (2\text{-}8)$$

The important characteristic of this transfer function is that the coefficients, $a_1, a_2, \ldots, a_7$ and $b_1, b_2, \ldots, b_7$

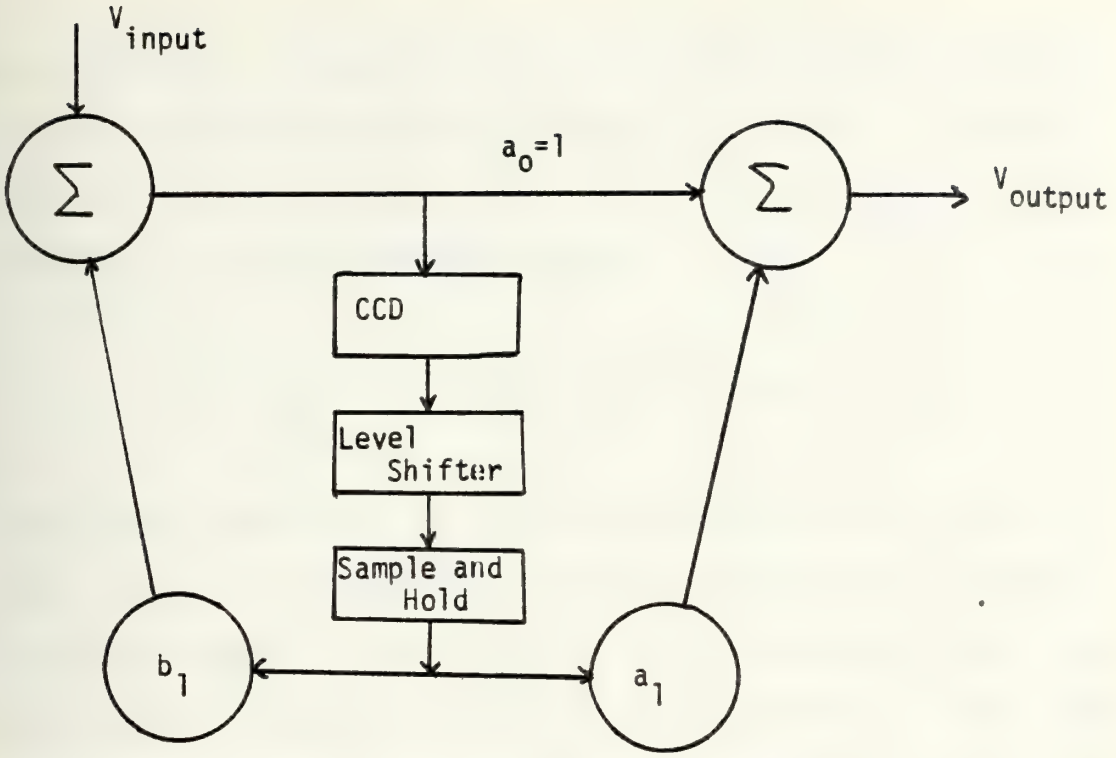


Figure 2.3. Experimental set-up for first order recursive filter

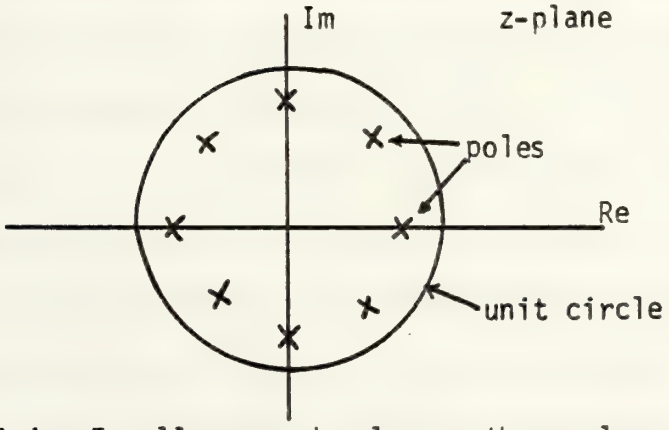


Figure 2.4. Equally spaced poles on the z-plane. An example of the eight to first order translation possible with serial-in, serial-out delay elements.

are all zero. Clearly, this is an eighth order filter. It has eight poles and eight zeros. In order to find the position of these poles and zeros on the z-plane, it is necessary to solve the equations

$$z^8 = \frac{-a_8}{a_o} \qquad \text{and} \qquad z^8 = -b_8 \tag{2-9}$$

Such a solution involves the eighth root of some real number (the coefficients, of course, cannot have nonzero imaginary parts). Now, poles lie at the various positions of the eight eighth roots of $-b_8$. In accordance with the mathematics of complex variables, each of these roots will have equal absolute values or moduli. The angles will, in general, be different. The mathematics of complex variables also tells us that the angles between the adjacent roots will be equal. That is, the poles (and zeros) will be EQUALLY SPACED angularly throughout the z plane. Figure 2-4 shows an example of such a plot.

Kuo [(11)] has shown that, in the same manner as is done in analog system transfer functions, the response of the sampled analog filter in the frequency domain can be obtained by graphical means. This method consists of picking a point on the unit circle and determining the length of the line segments drawn to each singular point on the plane. The lengths of these line segments are arranged so that the following quotient is obtained:

$$H = \frac{\text{distances from point to zeros}}{\text{distances from point to poles}} \qquad (2\text{-}10)$$

and H is the absolute value of the frequency response of the filter at the frequency of interest (determined by the point on the unit circle).

Using this method, we can investigate the frequency response of our eight pole-eight zero filter. From dc to one eighth of the sampling frequency, $f_s/8$, the distances to the singular points are found and the expression in Equation (2-10) is evaluated. In this way a point-by-point computation of the transfer function is made. When we begin to consider values of f larger than $f_s/8$, however, we notice that the singular points are in precisely the same relative location as they were for appropriate frequencies between dc and $f_s/8$. Thus, it is clear that the distances will be exactly equal, and further, the expression for H will be exactly equal for frequencies f, $f+f_s/8$, $f + 2f_s/8$, ..., $f + nf_s/8$.

This observation, in turn, is the very nature of periodicity. Thus, the transfer function values are periodic with respect to $f_s/8$.

To backtrack a little, it is well known that a sampled-analog filter is periodic with respect to f_s and is symmetrical about $f_s/2$. What we have done is to determine a whole class of recursive sampled analog filters which are not only

periodic with respect to f_s, but also periodic with respect to $f_s/8$. It turns out that this also implies that our new filter is symmetric with respect to $f_s/16$.

The use we are going to make of this discovery was foreshadowed by the first paragraph in this section. This filter, although nominally eighth order, can be considered a first order recursive sampled analog filter with an adjusted sampling frequency of $f_s/8$. Of course, certain adjustments must be made in the coefficients as well. For instance, in our set-up shown in Figure 2.3, if $b_8 = 0.7$, then the poles of the transfer function must lie on the $\sqrt[8]{0.7}$ circle on the z-plane. However, when the translation to first order is made, the pole formed by b_8 will lie on the real axis (since there is only one pole) and its modulus will be 0.7.

It should be understood, that for clarity, we have used in this discussion an eighth order filter. The theory and proof are equally good for any order filter as long as the condition that the middle coefficients are zero is met. An analytical proof is given in Appendix A.

The improvement in the actual design procedure by using this translation is perhaps better appreciated if one visualizes the design engineer at his bench attempting to find the best coefficient for a given ideal transfer function. Of course, there may be other methods (such as with a computer) which he could use, but one time-honored method is actual laboratory experiments. Unless he is unusually

bright, he cannot project the changes in the amplitude of the transfer function based on a small change of a coefficient in an eighth order filter. On the other hand, we have seen how simple it is to project changes in the transfer function of a first order filter by virtue of changes in the single coefficient, b_1. Thus the engineer is provided with an intuitive aid to design ... without resort to the digital computer.

A more fundamental reason for design in this manner concerns the construction of comb filters. Comb filters will be defined here as those electronic devices which exhibit a periodic transfer function, and which are little affected by the various nonideal circumstances which inevitably appear in any physical system. For instance, our theory has depended on the ability to sample a waveform instantaneously and store a number which corresponds to the value of the waveform at that instant. In addition, we have assumed that we can restore the original signal by using an ideal low pass filter, when, in fact, most of the restoration of signals is done with a sample-and-hold device, which is NOT ideal. Finally, we have assumed that our delay line is free from attenuation and distortion. In a charge coupled device, we always have a certain amount of charge left behind during each transfer from gate to gate. The charge left behind adds to the succeeding packet of charge and causes amplitude distortion. These factors, plus other instrumentation

realities, prevent us from ever producing a truly periodic transfer function. The sample-and-hold effect causes a sin x/x envelope to be superimposed in a multiplicative way on the sampled-analog filter transfer function. The non-zero sampling time also causes an additional six x/x envelope to be superimposed multiplicatively. The order of these two effects is not the same, however. The six x/x caused by the sample-and-hold is far more restrictive since the first zero of the sin x/x function occurs at $f = f_s$. The first zero of the other effect occurs at $f = 1/\tau$ where τ is the sampling time width, usually quite small.

Thus, in designing a comb filter by conventional means, the use of sample-and-hold circuitry inherently causes a rather rapid fall off of amplitude response. That is to say, if there is only one maximum of the transfer function between dc and f_s, then the second maximum, located between f_s and $2f_s$, is much smaller than the first. The ideal comb filter, of course, has equal maxima throughout the spectrum.

Now consider the use of the method explained above. There are as many maxima between dc and f_s as there are delay elements in the delay line, which is, of course, equal to the true order of the filter. In our case, this number was eight. Thus eight maxima appeared in the region between dc and f_s. The sample-and-hold circuitry still caused the sin x/ x rolloff of amplitude response, but its effect was negligible for most applications for the first two maxima. The real payoff comes when additional delays are built into

the system. One such circuit which we have used is the RETICON SAD (serial-analog delay device). This circuit has 96 delay elements and the sampled analog filter transfer function (using this device) has, in theory and practice, 96 maxima between dc and f_s. Clearly, the first twenty or so are not affected adversely by the sin x/ x factor.

Thus, it is possible, using these techniques, to design an almost ideal comb filter while, at the same time, offering to the design engineer the enormous advantage of being able to treat his filter as if it were a first order filter with the accruing simplicity of design. Figure 2.5 shows the first four "teeth" or periods of a comb filter designed by conventional means, corrected for the effect of the sample-and-hold circuitry. Figure 2.5b shows four "teeth" or periods of a comb filter designed by the aforementioned technique using a delay of eight stages. Finally, Figure 2.5c shows the same four "teeth" using a 96 stage delay device. The comparison between the three is clear. The latter exhibits the best performance for periodicity.

SECTION 2.2 EXPERIMENTAL RESULTS AND THEORETICAL CALCULATIONS OF REPRESENTATIVE FIRST ORDER RECURSIVE COMB FILTERS

2.2.1 With the theory of Section 2.1 in mind, the next step was to determine whether adequate first order filters could be implemented using the charge coupled device as the delay element. The CCD selected for the experiment was built by TRW SYSTEMS and consisted of an eight bit serial-in,

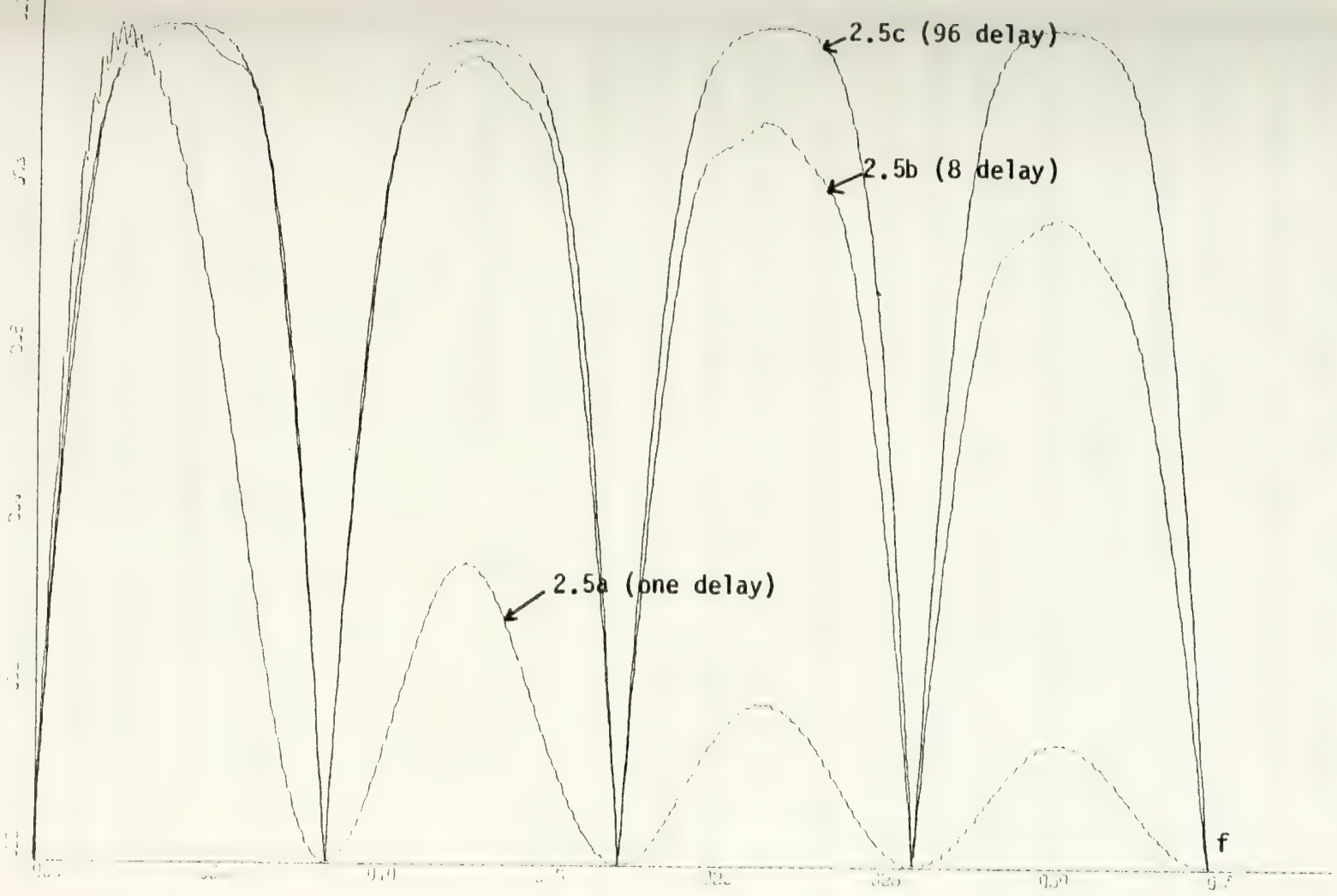


Figure 2.5. Transfer functions of one (2.5a) delay, eight delay (2.5b), and 96 delay (2.5c) first order sampled analog filters. Notice that the 96 delay device provides the best uniformity among "teeth". All curves corrected for sample-and-hold effects

serial-out delay register. The cross section of the device is shown in Figure 2.6. This is the same device as was used for the investigation of the theory of CCD operation. The curves of operation derived in Chapter one were utilized here to find an appropriate biasing point. To choose the optimum point, we desired to have the widest possible linear range for both input and output. Such a point was found to occur whenever the first two input gates (G_1 and G_2) were biased low (-30 v), the input impressed upon gate #3, while using the voltage on gate #4 as a parameter. The actual operating curve is shown in Figure 2.7. The linear ranges are easily seen. The steep slope on the left is the so-called barrier mode of operation which was discussed in Chpater one. The second region is the region of "scuppering", "sloshing", or "potential equilibration". The actual quiescent point used was

$$V_3 = -18.7 \text{ v}$$

$$V_4 = -12.0 \text{ v}$$

which provided a linear range of approximately 3.7 v.

Figure 2.3 shows the experimental set-up. For $a_o = 1$, a direct connection was made from the input summer to the output summer. For a_1 and b_1, potentiometers were connected into the circuit after the sample-and-hold and level shifter. The full value of these potentiometers was 20 KΩ. The value for a given coefficient was determined by considering the

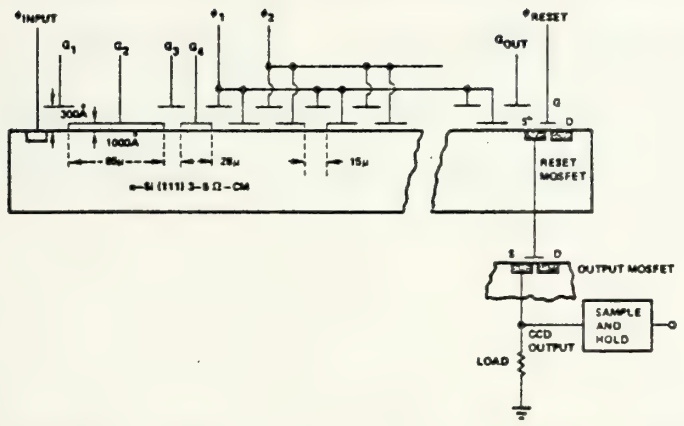


Figure 2.6. Cross section of CCD.

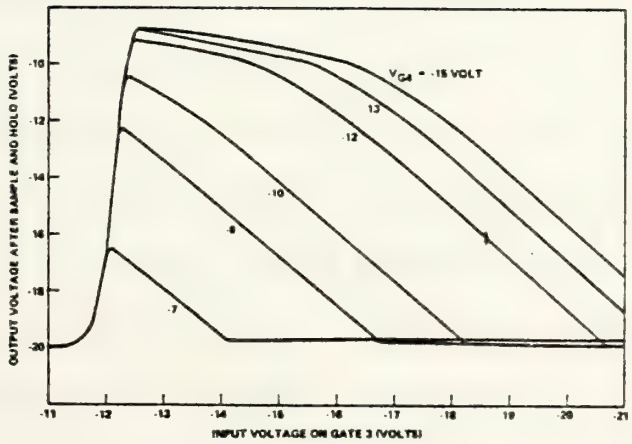


Figure 2.7. Operating curves for CCD selected for use in filter design.

gain of the CCD and the input impedance of the summers. The value of the potentiometers was, then, not obviously related to the value of the coefficients in the transfer function. Specifically, if the input impedance of the summer could be neglected, then the value of the potentiometer, P, was related to the coefficient, a, by

$$P = a/K$$

where K is the overall gain of the CCD, sample-and-hold device, and level shifter.

Unfortunately, we found, for first order experiments that the operational amplifier summers did not have large enough input impedances to justify neglecting them. Thus, we needed a correction factor for the summers. The applicable circuit diagram is shown in Figure 2.8. The input to the circuit (output circuitry of the level shifter) is KV_{in} and the desired output is aV_{in}. Using the usual virtual ground assumption for operational amplifiers, we can calculate our desired value of the potentiometer which yields aV_{in}. We call this value P. The total impedance as seen from the output of the level shifter is

$$Z_{tot} = 20K - P + \frac{PZ_{in}}{P + Z_{in}} \qquad (2\text{-}11)$$

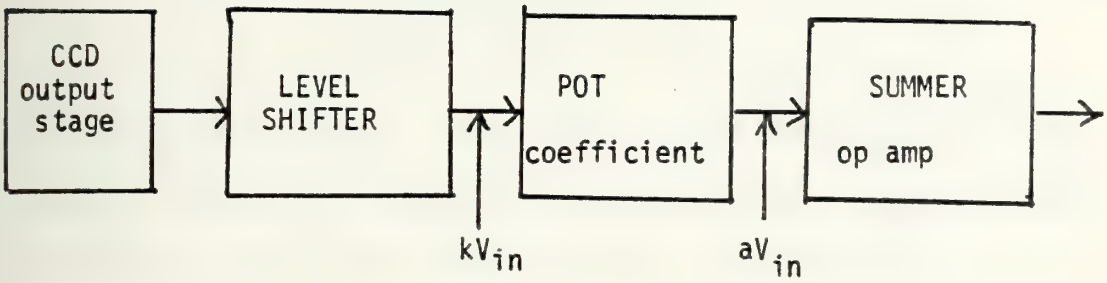


Figure 2.8. Block diagram of implementation of feedback and feed forward coefficients.

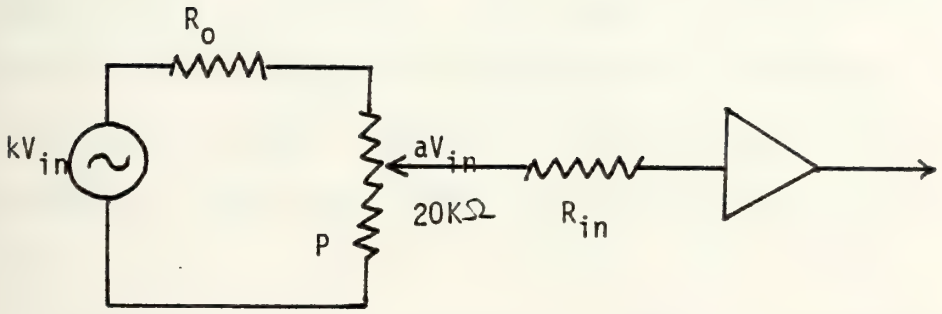


Figure 2.9 . Voltage divider circuit for coefficient implementation. Correction factors were calculated which included the effects on R_{in}. R_o was very small.

This becomes a voltage divider network as shown in Figure 2.9. Thus our voltage divider transfer function becomes

$$aV_{in} = \frac{k\ V_{in}\ P\ Z_{in}}{P\ 20K + 20K\ Z_{in} - P^2} \qquad (2\text{-}12)$$

In the above equation, all the quantities except P are known. Thus, in order to assist us in our experiments, a correction curve was generated which related "a" to "P" for all appropriate values of a ($a < 1$).

In order to have a coherent set of experiments, it was decided to utilize both the standard z and bilinear z transforms to generate first order filters. In addition, we required that a number of integrator and canceller types of filter be investigated. The integrator type filter is defined as having narrow pass bands (high Q) with wide stop bands where attenuation is high. The canceller is the inverse of the integrator. The canceller has narrow stopbands (high Q where Q for stopbands is defined as the Q for a filter whose transfer function is 1 - H(f)) with wide passbands. These two types of filters are shown in Figure 2.10.

Naturally, it was desired to look at low pass and high pass type filters as well. Using these designations, it is possible to consider a low pass, integrator type, filter designed by the standard z-transform or any other combination of these three categories as shown in Table 2.2.

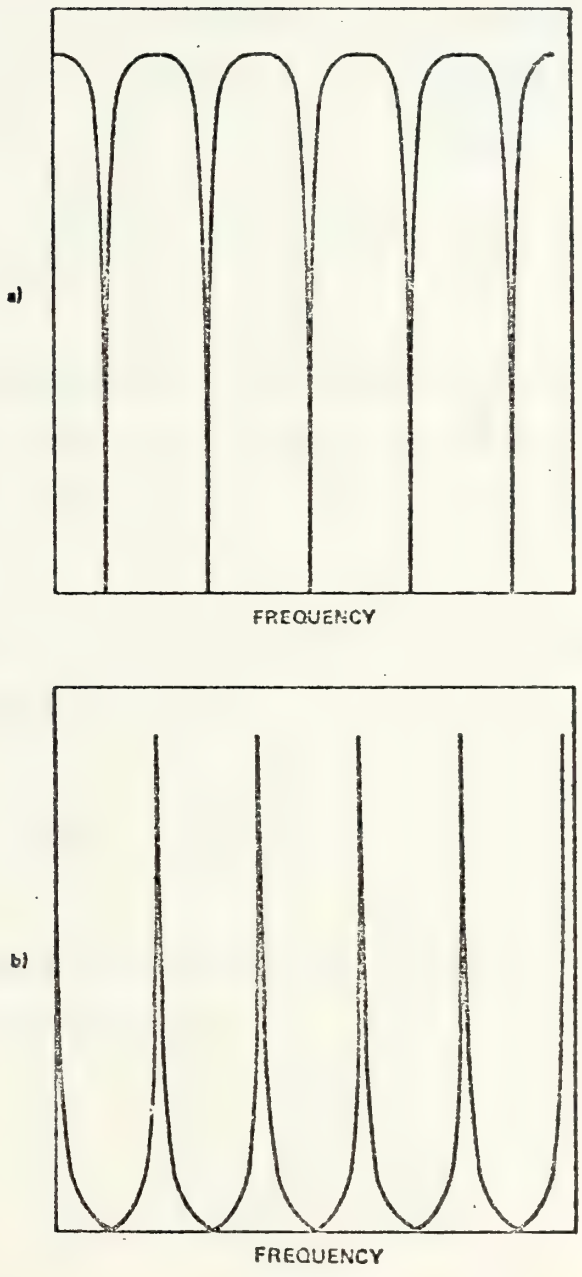


Figure 2.10 Frequency characteristics of two general types of comb filters - - - bandstop, canceller type; and bandpass, integrator type

PASS	TYPE	DESIGN METHOD
LOW	INTEGRATOR	STANDARD Z
HIGH	CANCELLER	BILINEAR Z
		OTHER

Table 2.2: Categories of first order filters

In order to have general expressions for the coefficients (so the coefficient would not have to be derived every time we needed a change in its value), we made the following calculations:

For the standard z transform, we assume a form for the low pass filter to be

$$H(s) = \frac{\omega_x}{s + \omega_x} \qquad (2\text{-}13)$$

and find the standard z transform of this H(s). The result is (except for a gain factor):

$$H(z) = \frac{1}{1 - e^{-\omega_x T} z^{-1}} \qquad (2\text{-}14)$$

Such a filter is an integrator if the cutoff frequency is less than 0.25 times the sampling frequency, and a canceller if the cutoff frequency is greater than 0.25 times the sampling frequency. These statements are the logical results of our definition of integrator and canceller. When the

cutoff frequency is exactly $0.25f_s$, the passband is precisely as wide as the stopband and neither an integrator nor a canceller can be achieved. On the other hand, it is well to keep in mind that, by using the standard z transform, we may not get any cutoff frequency at all.

The requirement for the integrator is that the cutoff frequency be less than $0.25f_s$ $(0.25/T)$. This translates to a requirement that $b_1 = -e^{-\omega_x T}$ be less than -0.268 (this calculation is found in Appendix B). A quick look at exponential tables tells us that, in order to have a low pass integrator, we must have $f_x T$ less than 0.21. This last statement tells us something about the relationship between the s-plane analog filter and the sampled analog filter of the z-plane. The ω_x is the radial cutoff frequency of the analog filter. If we desire a narrow band low pass filter in the z-plane, we must choose a sampling frequency, f_s, such that $f_x < 0.21f_s$ or a sampling frequency considerably higher than the analog cutoff frequency. Suppose that the analog cutoff frequency was 10 kHz. Then in order to design a low pass integrator, we must require a sampling frequency of at least

$$f_s = \frac{10^4}{0.21} = 47.6 \text{ kHz} \qquad (2\text{-}15)$$

Failure to do so may yield a wide band low pass filter, or, worse, a quasi all pass filter (one with no cutoff frequency).

Naturally, to achieve a low pass canceller, we need to have $f_x/f_s > 0.21$. We also have a lower limit below which we get no cutoff frequency. This is found by determining at what sampling frequency the cutoff frequency no longer exists. This calculation is found in Appendix D. The critical value for b_1 such that if b_1 takes on any larger value, no cutoff frequency exists is -0.172. This translates, via exponential tables to

$$- \omega_x T = - 1.76 \qquad (2\text{-}16)$$

which further reduces to the requirement that

$$f_x/f_s < 0.28 \qquad (2\text{-}17)$$

Thus, we also have determined what the appropriate sampling frequency will be in order to design a low pass canceller:

$$0.21 < f_x/f_s < 0.28 \qquad (2\text{-}18)$$

Similar calculations can be made for high pass filters. Except for the sign of b_1 (it becomes positive since the poles must lie on the negative real axis of the z-plane), the values are the same:

INTEGRATOR: $0 < f_x/f_s < 0.21$

CANCELLER: $0.21 < f_x/f_s < 0.28$

One other possibility exists in using these standard z transform filters. If we simply subtract the transfer function (normalized to one) from unity for all frequencies, we can achieve the inverse of each filter type with different coefficients. For instance, if we subtract the transfer function for the low pass filter

$$H(z) = \frac{1}{1 - b_1 z^{-1}} \qquad (2\text{-}19)$$

from one, we get

$$H(z) = \frac{-b_1 z^{-1}}{1 - b_1 z^{-1}} \qquad (2\text{-}20)$$

which is a high pass filter. If the original filter was an integrator type, then the latter filter is a canceller. This translation produces the following results

Start with	TRANSLATION (1-H(z))	Result
Low Pass integrator		High pass canceller
Low pass canceller		High pass integrator
High pass integrator		Low pass canceller
High pass canceller		Low pass integrator

One remark about these filters is worth making here. We observe that for all these filters, only one term appears in the numerator, either a_o or $a_1 z^{-1}$. In either case, it is clear that no true zero will ever appear in the transfer function for real frequencies. Looking at the pole-zero plot, this means that a zero will never appear on the unit circle. If we are seeking a large passband-to-stopband attenuation ratio, it appears that the only means to achieve our goal is by seeking the optimum value of b_1. Since b_1 is a function of the ratio of the analog cutoff frequency to the sampling frequency, this means that in searching for optimum values of b_1, we are actually searching for the optimum ratio of analog cutoff frequency to sampling frequency. To be more explicit about this procedure, let us consider an example: If we have transformed an analog filter whose cutoff frequency was f_c using a sampling frequency of f_s, we have obtained a b_1 of $\pm e^{-\omega_c T}$. Now, if we observe the frequency response of this filter and decide to shift it by changing b_1 slightly, what we have actually done (assuming our sampling frequency remains constant) is to change the original analog filter from which our sampled analog filter was derived.

Table 2.3 shows the generalized formulas for the coefficients for a filter derived by the standard z transform.

Figures 2.11 through 2.14 show the experimental data and the theoretical shape of the transfer functions for four

TABLE 2-3

Design Summary of First Order Filters
Using the Standard z Transform

FILTER TYPE	a_o	a_1	b_1
Low Pass Integrator			
$\frac{f_c}{f_s} < 1/4$	1	0	$-e^{-\omega_c T}$
Low Pass Canceller	0	$-e^{-\omega_c T}$	$e^{-\omega_c T}$
$\frac{f_c}{f_s} > 1/4$			
High Pass Canceller	0	$-e^{-\omega_c T}$	$-e^{-\omega_c T}$
$\frac{f_c}{f_s} < 1/4$			
High Pass Integrator	1	0	$e^{-\omega_c T}$
$\frac{f_c}{f_x} > 1/4$			

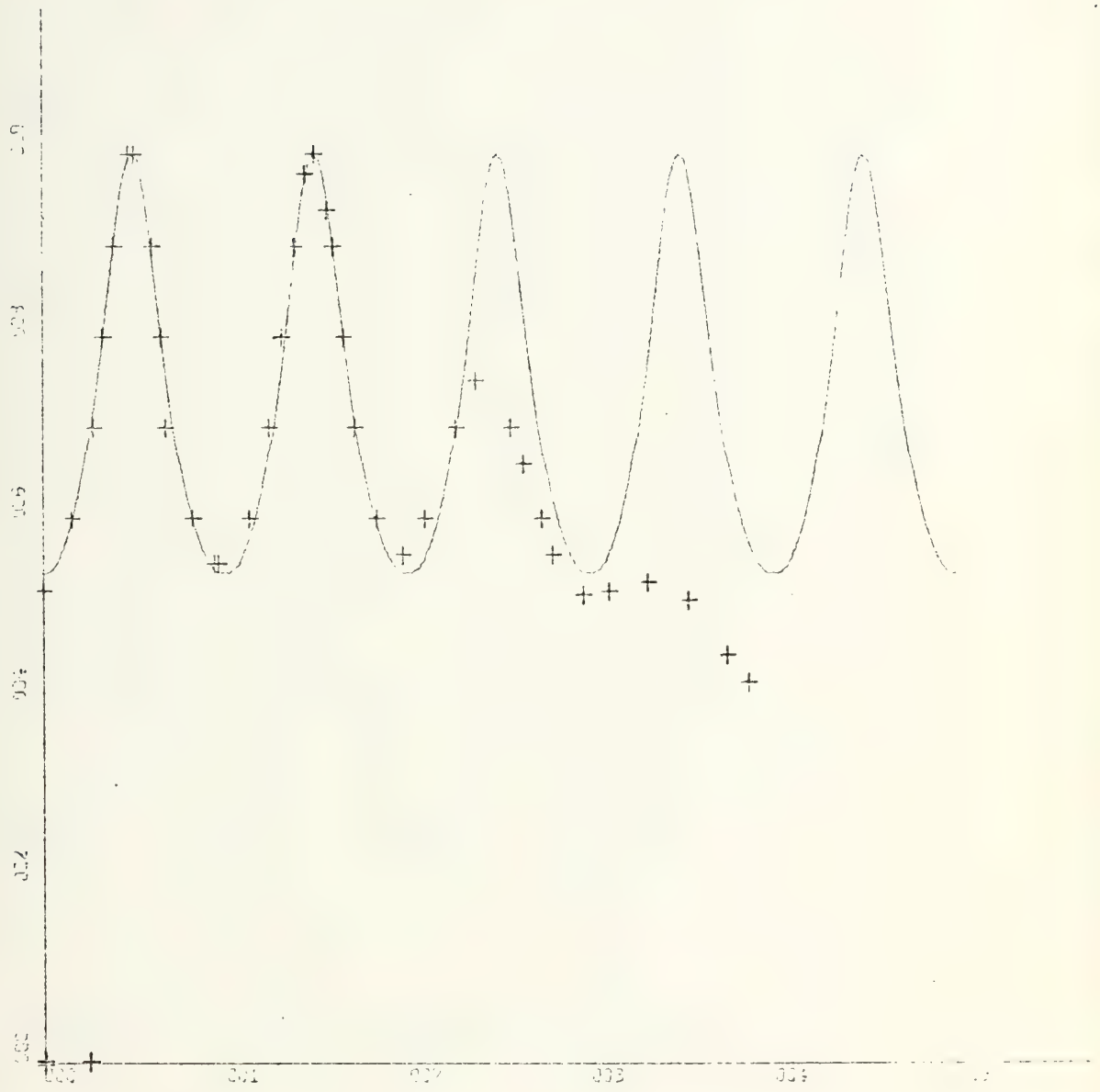


Figure 2.11. Theoretical and Experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 0$, $b_1 = 0.3$ Experimental data is shown by a (+).

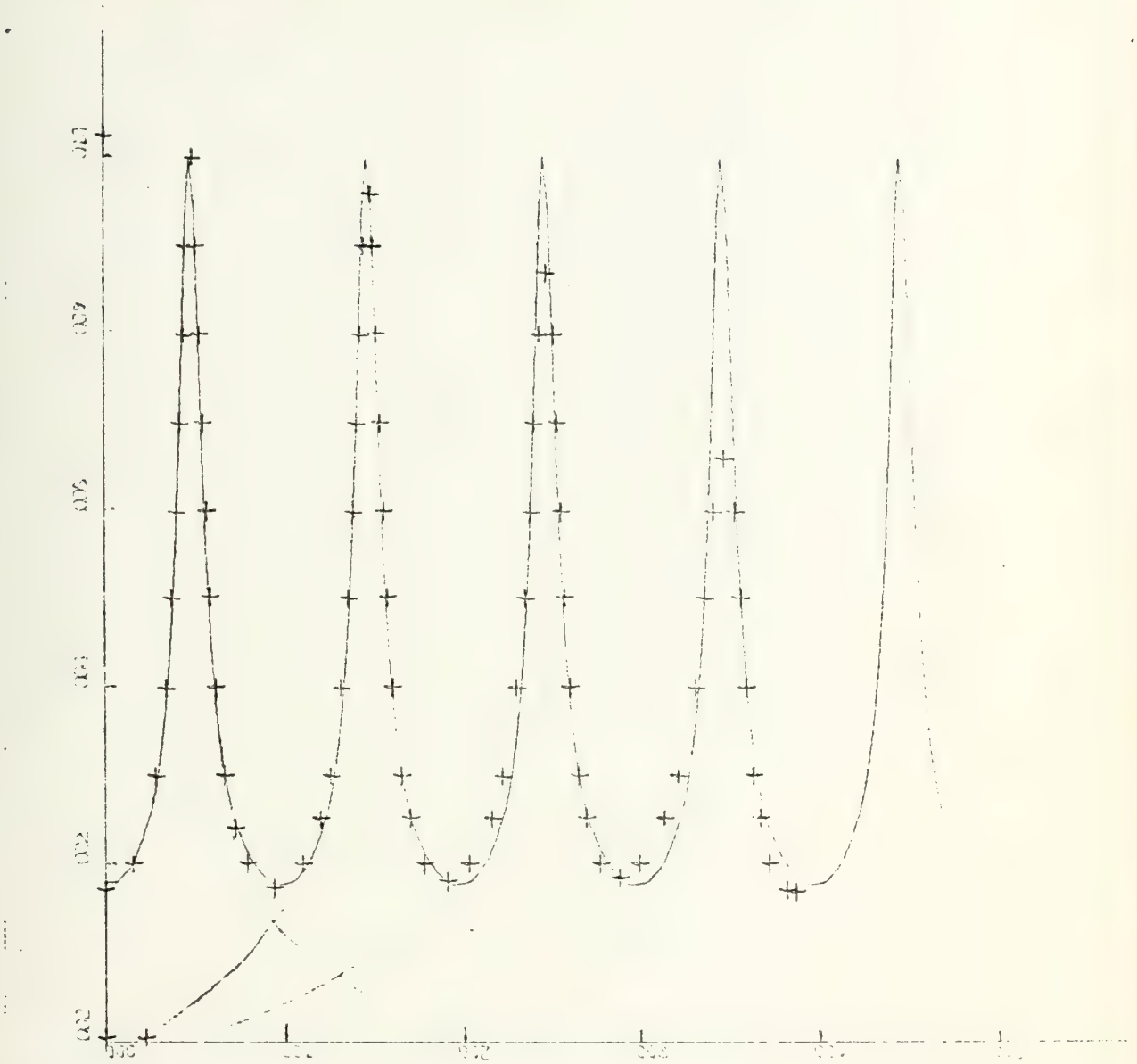


Figure 2.12 Theoretical and Experimental transfer function for a first order recursive filter, $a_o = 1$, $a_1 = 0$, $b_1 = 0.7$ Experimental data is shown by a (+).

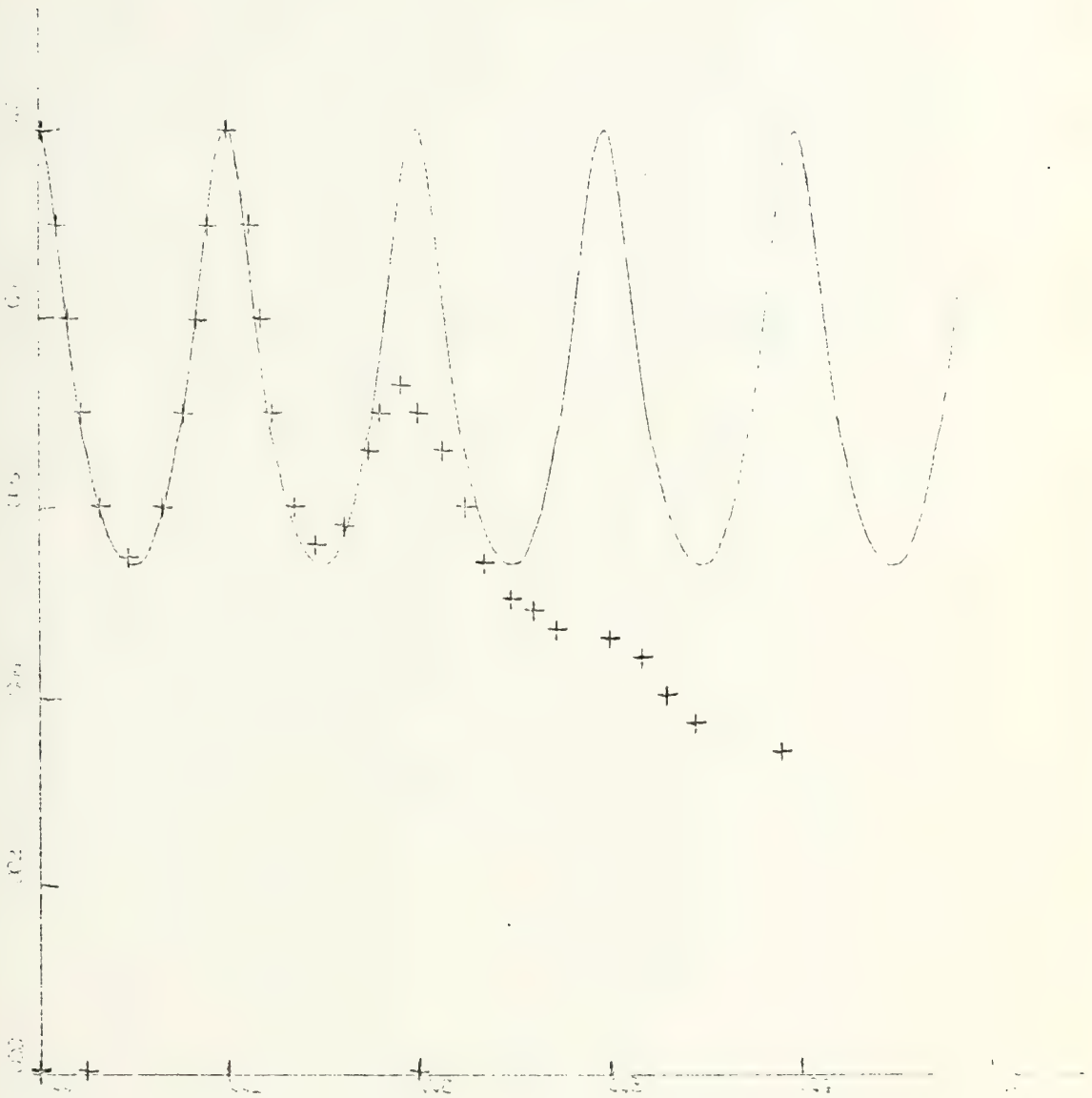


Figure 2.13 Theoretical and Experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 0$, $b_1 = -0.3$ Experimental data is shown by a (+).

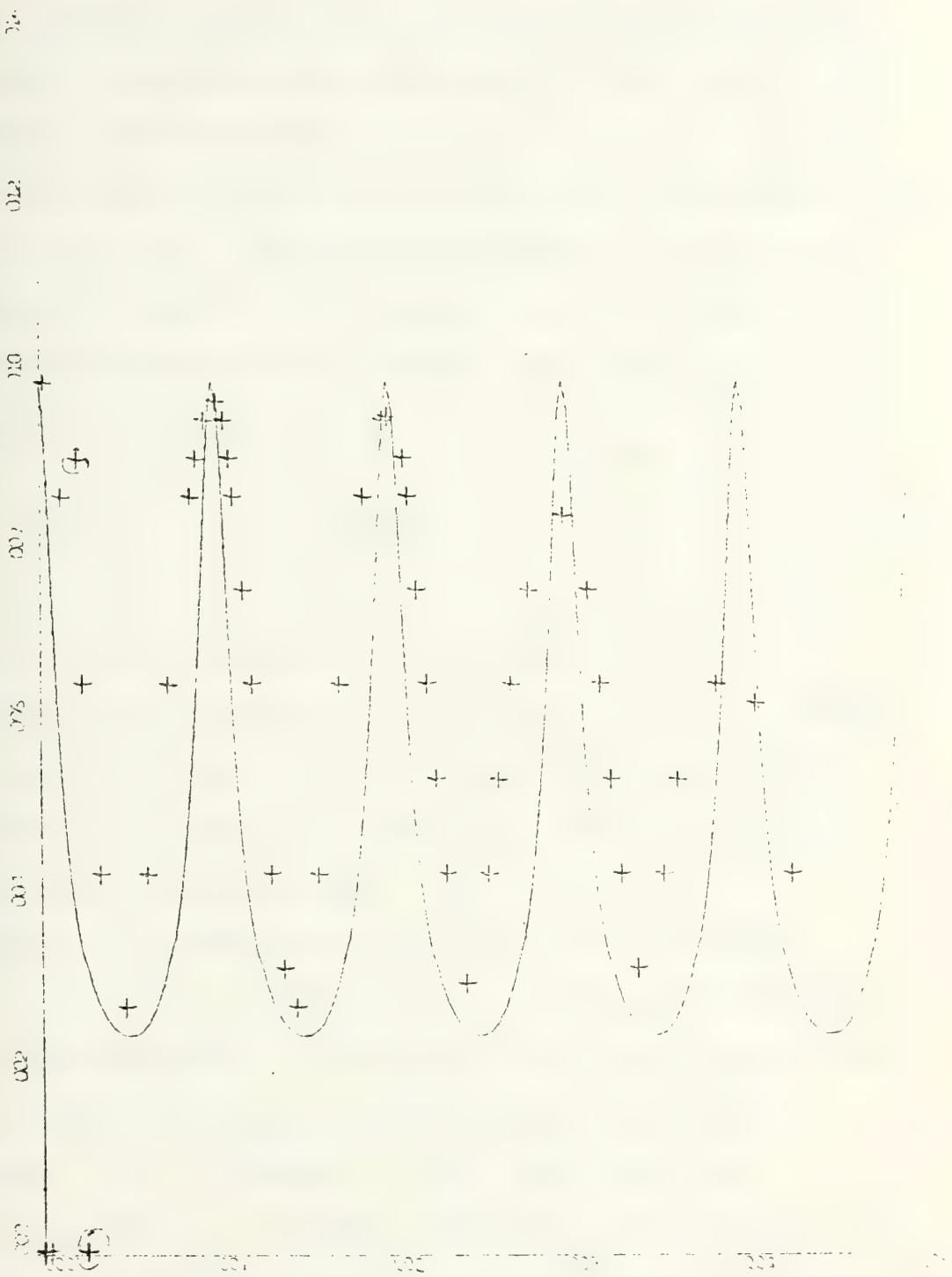


Figure 2.14. Theoretical and Experimental transfer function for a first order recursive filter, $a_o = 1$, $a_1 = 0$, $b_1 = -0.6$ Experimental data is shown by a (+).

filters. The first filter was designed as a high pass integrator. The feedback coefficient is fairly small, 0.3, which corresponds to a f_x/f_s ratio of 0.19. From the discussion above, it can be confirmed that this ratio falls into the range of a high pass integrator. Nevertheless, the ratio is "close" to the limit, 0.21, and therefore, we would expect that the filter would have a poor quality factor, where Q is defined by

$$Q = \frac{f_{center}}{BW} \tag{2-21}$$

where BW is the bandwidth of the filter.

Increasing the feedback coefficient, b_1, to 0.7 enables us to design and test a filter with considerably higher selectivity (a higher Q). Figure 2.12 shows the results of the theory and experiment.

Figure 2.13 shows our results for small feedback ($b_1 = -0.3$) and a low pass filter. The shape of the frequency characteristic is the same as for the high pass case with the same absolute value of feedback coefficient. Therefore, it is, of course, a low pass integrator.

Finally, Figure 2.14 shows the results for large negative feedback. Apparently, during the experiment, the feedback coefficient was less than had been expected since the experimental and theoretical curves do not correspond so closely as the other three cases.

We observe that, as expected, none of the four cases show a true zero either experimentally or theoretically. This is in accordance with our predictions. Agreement between theory and practice was good except for the condition of large negative feedback. These curves all show a linear plot, so that the maximum attenuation ratio between peak and null occurs for the higher feedback case and is approximately 12 db. The same ratio for the condition of low feedback ($b_1 = 0.3$) is less than 6 db. Clearly filters should be able to achieve better passband-to-stopband ratios than this. A search for better filters leads us to the next section.

2.2.2 The Bilinear Z Transform

We start with the same analog filter as before and apply the bilinear z transform directly to the transfer function.

$$H(s) = \frac{\omega_x}{s + \omega_x} \tag{2-22}$$

Applying the bilinear z transformation:

$$s = \frac{2}{T}\frac{z - 1}{z + 1} \tag{2-23}$$

we get

$$H(z) = \frac{\omega_x T + \omega_x T z^{-1}}{(\omega_x T + 2)(1 + z^{-1}\frac{\omega_x T - 2}{\omega_x T + 2})} \tag{2-24}$$

Now using the same criterion for integrator and canceller type filters, we must find the applicable ratios of f_x/f_s and appropriate values of a_o, a_1, and b_1 to use.

One good feature of the bilinear z transform is that it produces both nonzero a_o and a_1. In this way, we can be sure that a theoretical zero will occur in the transfer function. Of course, we will not expect an infinite passband-to-stopband response ratio, but the ratio should be better than in the standard z transform case. We note that the value for a_o and a_1 are precisely equal in Equation (2-24). On the z-plane this corresponds to a zero on the unit circle at z = -1.0. Since the transfer function will be continuous, we know that there will be a cutoff frequency ... there is no uncertainty as there is when using the standard z transform. Of course, the cutoff frequency will, in general, not be the same cutoff frequency as existed in the analog filter since the two functions of frequency are different. But a cutoff frequency is certain to exist.

Requiring that the cutoff frequency of the sampled analog filter be less than $0.25f_s$, we find the following relations:

A. For the low pass integrator, the constraints are

$$-1 < b_1 < 0 \qquad \text{and} \qquad 0 < f_x/f_s < 1/\pi$$

B. For the low pass canceller, the constraints are

$$0 < b_1 < 1 \qquad \text{and} \qquad 1/\pi < f_x/f_s < \infty$$

where the value of the ratio f_x/f_s rises without bound as b_1 tends toward 1. The physical explanation for this is that the cutoff frequency in the analog domain may be as large as desired, and still, there will be a cutoff frequency in the sampled analog domain for ANY sampling frequency. Therefore, suppose we have a filter with a large cutoff frequency, say 100 kHz. Now suppose we wish to apply the bilinear z transformation to this transfer function and use the resulting filter in a system where the sampling frequency is 20 kHz. Now this sort of thing simply could not be done using the standard z transform, since, as explained in the preceding section, the sampling frequency must be considerably larger than the analog cutoff frequency in order to achieve proper cutoff in the sampled analog domain. However, using the bilinear z transform, we can be assured that the cutoff frequency will exist, and it will be greater than $0.25f_s$ or 5 kHz and be between 5 kHz and 10 kHz for a canceller. The power of the bilinear z transform is emphasized by these computations.

As a corollary to the existence of the cutoff frequency, we can also infer the existence of any other value of the transfer function so long as the value existed in the analog

filter transfer function. This occurs because the bilinear z transform maps the entire jω axis of the s-plane onto one revolution of the unit circle in the z-plane. So, just as we always have cutoff, we also always have a point which is, say, 30 db down from the maximum, as long as such a point existed in the analog filter function.

In the high pass case, the transfer function derived by the bilinear z transform is almost the same as for low pass except the coefficient a_1 is negative. Thus, this implies the existence of a zero in the z-plane at z = 1 (or f = 0, dc). The coefficient, b_1, once again varies from -1 to 1 as follows:

A. For the high pass canceller, the constraints are

$-1 < b_1 < 0$ and $0 < f_x/f_s < 1/\pi$

B. For the high pass integrator, the constraints are

$0 < b_1 < 1$ and $\infty > f_x/f_s > 1/\pi$

Table 2.4 shows the generalized formulas for the coefficients for a filter derived by the bilinear z transformation.

Figures 2.15 through 2.28 show various filter transfer functions for various coefficients. They are tabulated here for easy reference:

Figure no.	A_o	A_1	B_1	Type
2.15	1	-1	0.9	High pass integrator (high Q)

TABLE 2-4

Design Summary of First Order Filters
Using the Bilinear z Transform

FILTER TYPE	a_o	a_1	b_1
Low Pass Integrator	1	1	$-b_1$
$\frac{f_c}{f_s} < 1/4$			
Low Pass Canceller	1	1	$+b_1$
$\frac{f_c}{f_s} > 1/4$			
High Pass Canceller	1	-1	$-b_1$
$\frac{f_c}{f_s} < 1/4$			
High Pass Integrator	1	-1	$+b_1$
$\frac{f_c}{f_s} > 1/4$			

where, in all cases, $b_1 = \frac{\omega_c T - 2}{\omega_c T + 2}$

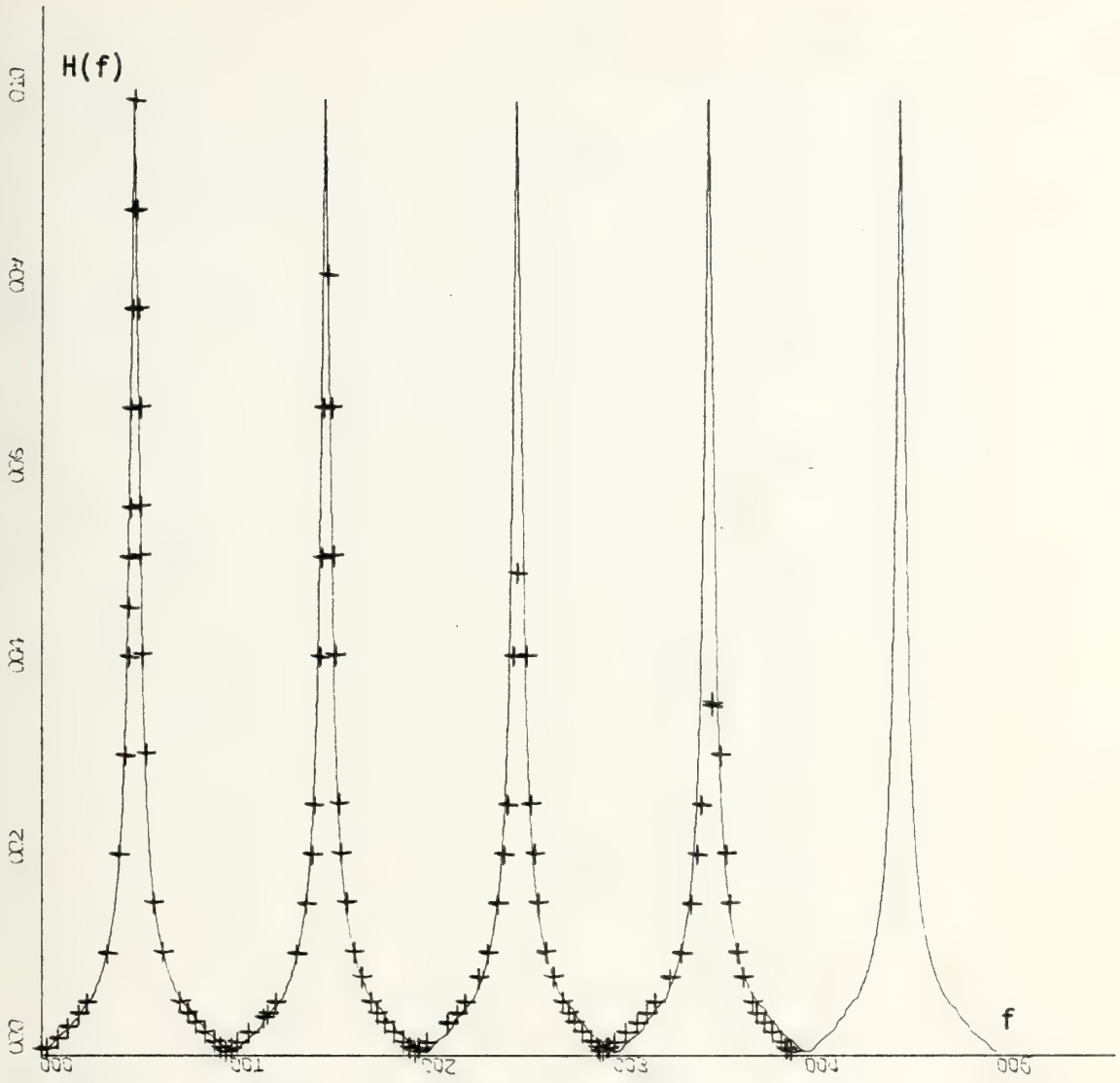


Figure 2.15. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = -1$, $b_1 = 0.9$ Experimental data are shown by a (+).

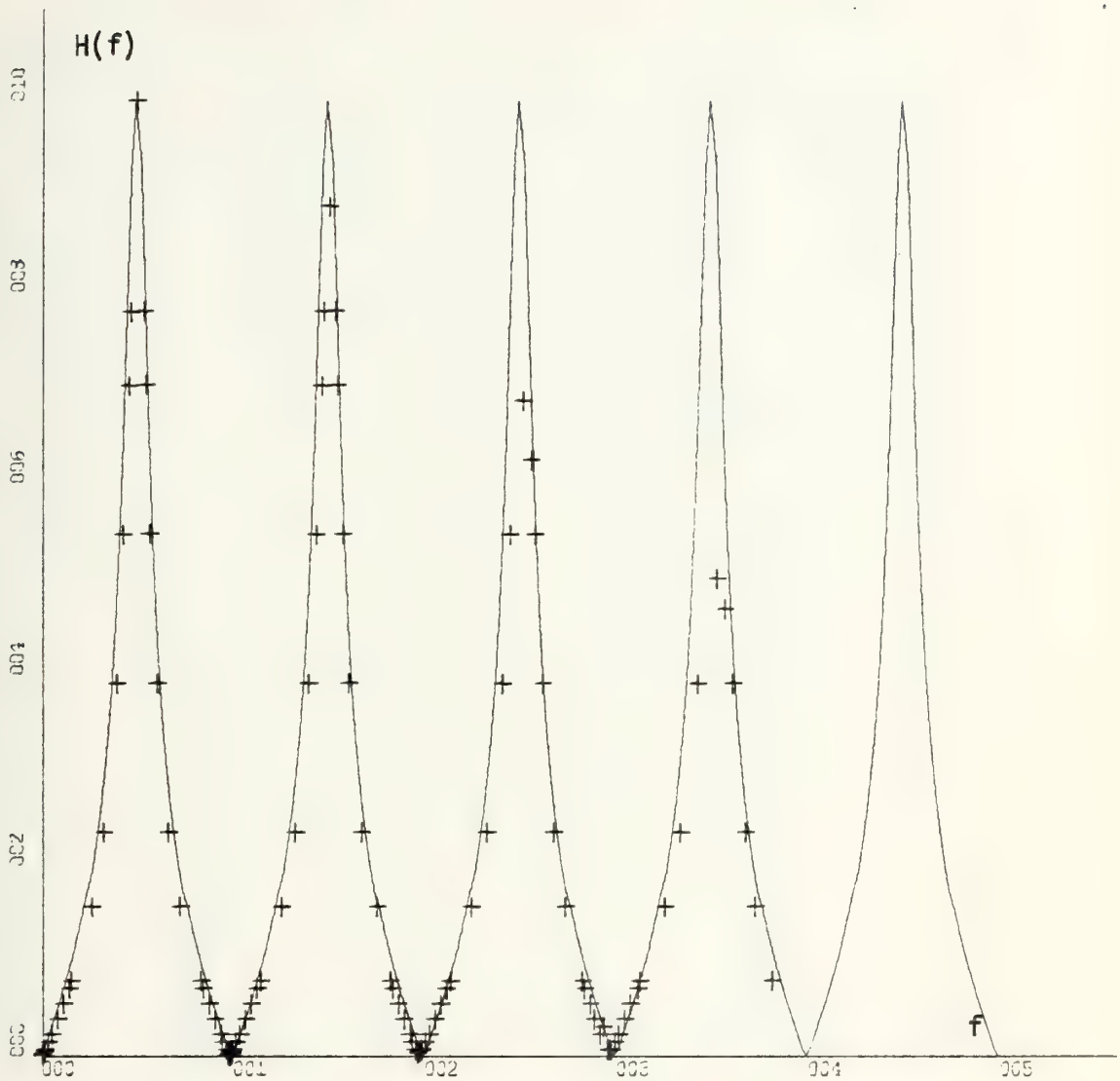


Figure 2.16. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = -1$, $b_1 = 0.7$ Experimental data are shown by a (+).

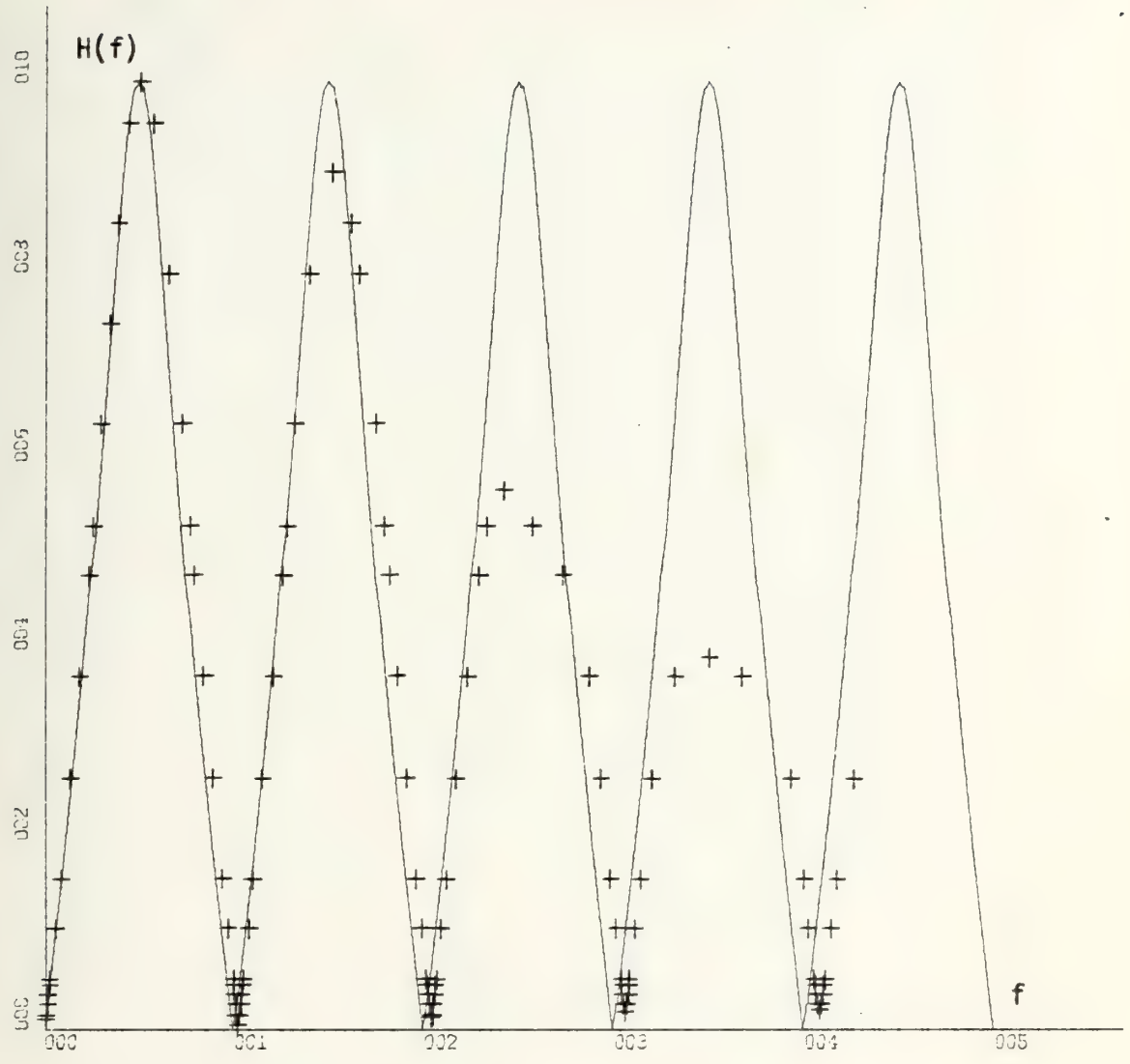


Figure 2.17. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = -1$, $b_1 = 0.3$ Experimental data are shown by a (+).

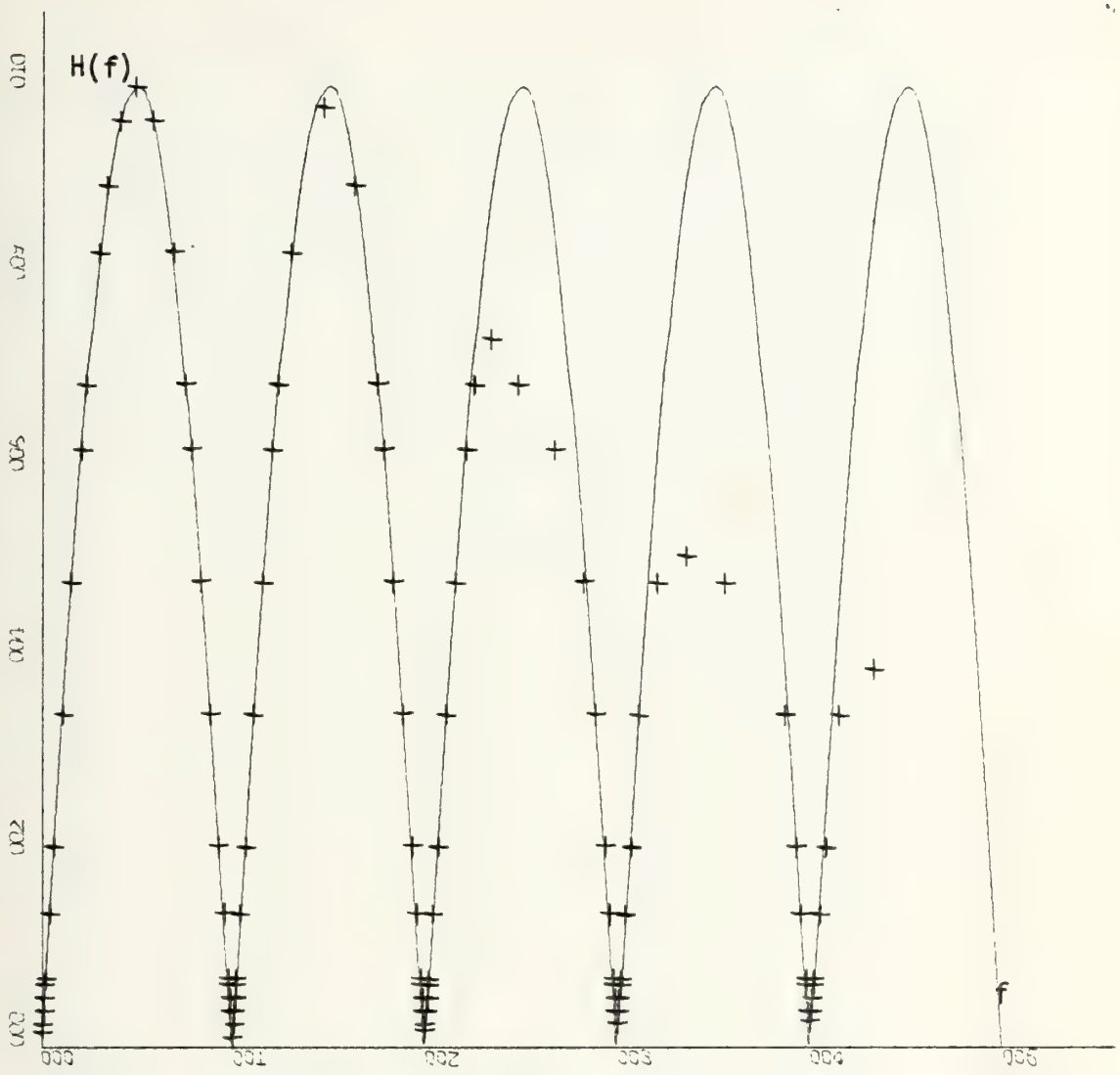


Figure 2.18. Theoretical and experimental transfer function for a first order transversal filter, $a_0 = 1$, $a_1 = -1$, $b_1 = 0$ Experimental data are shown by a (+).

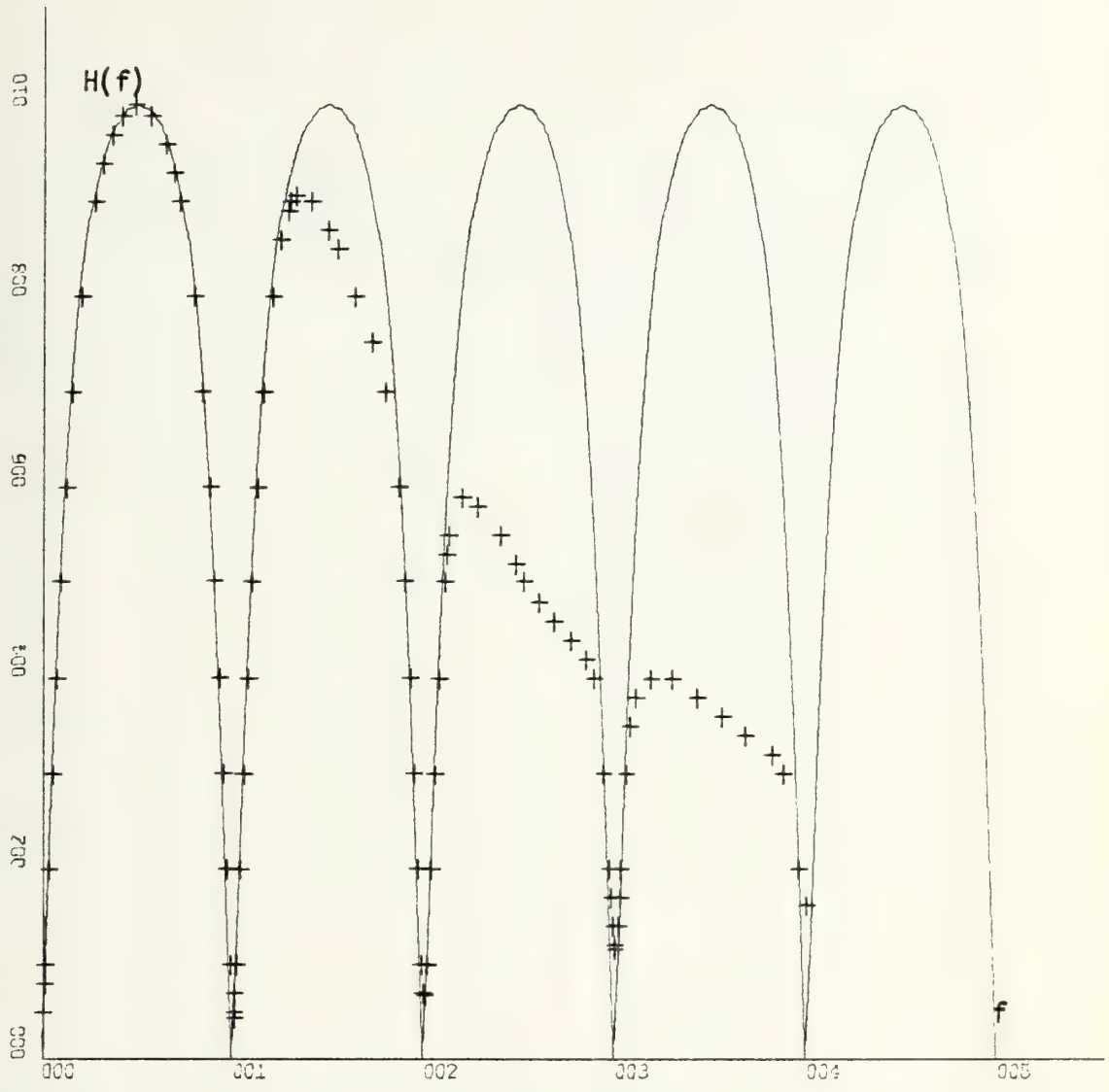


Figure 2.19. Theoretical and Experimental transfer function for a recursive first order filter, $a_0 = 1$, $a_1 = -1$, $b_1 = -0.3$ Experimental data are shown by a (+).

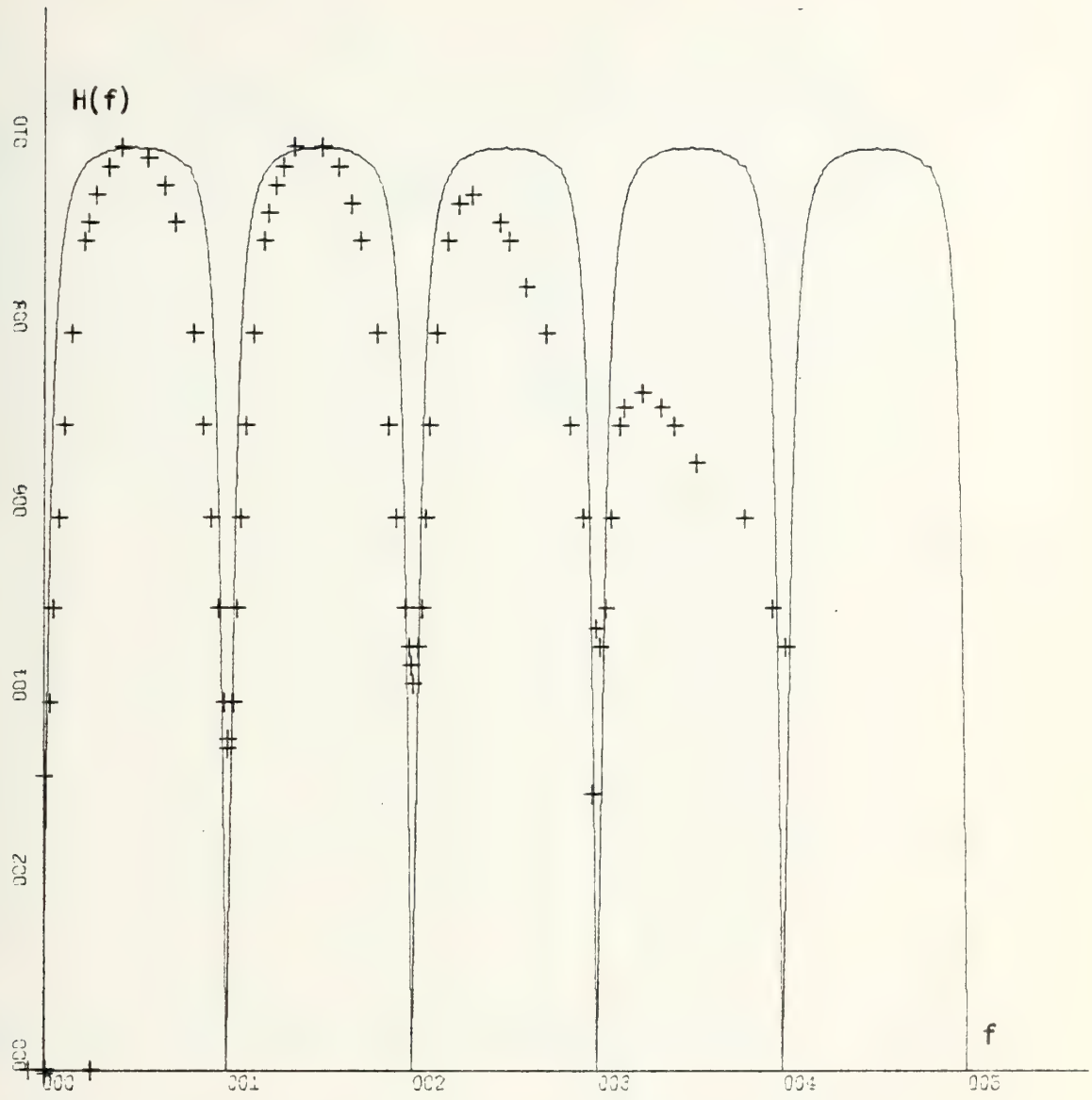


Figure 2.20. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = -1$, $b_1 = -0.7$ Experimental data are shown by a (+).

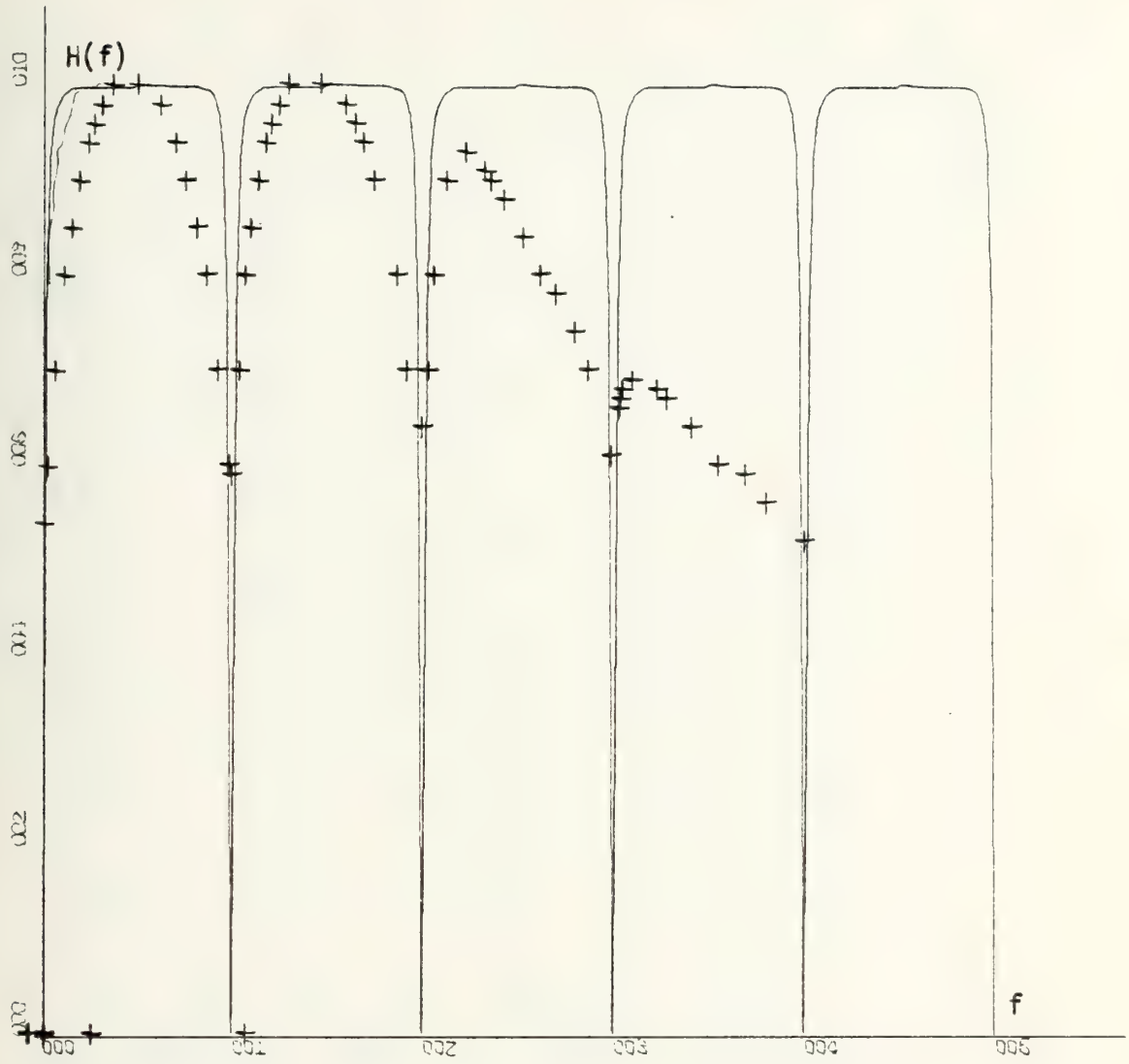


Figure 2.21. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = -1$, $b_1 = -0.9$ Experimental data are shown by a (+).

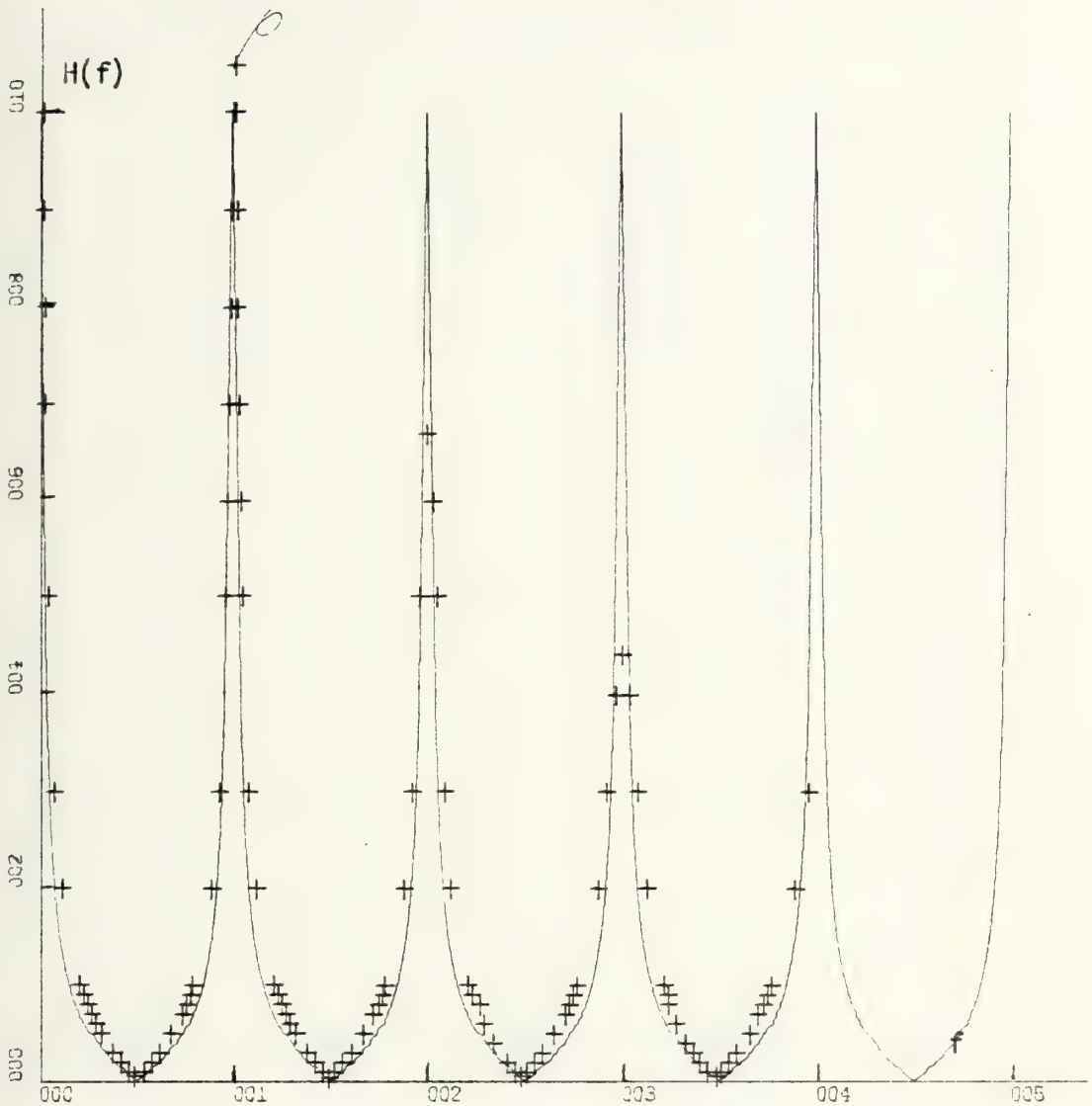


Figure 2.22. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = -0.9$ Experimental data are shown by a (+).

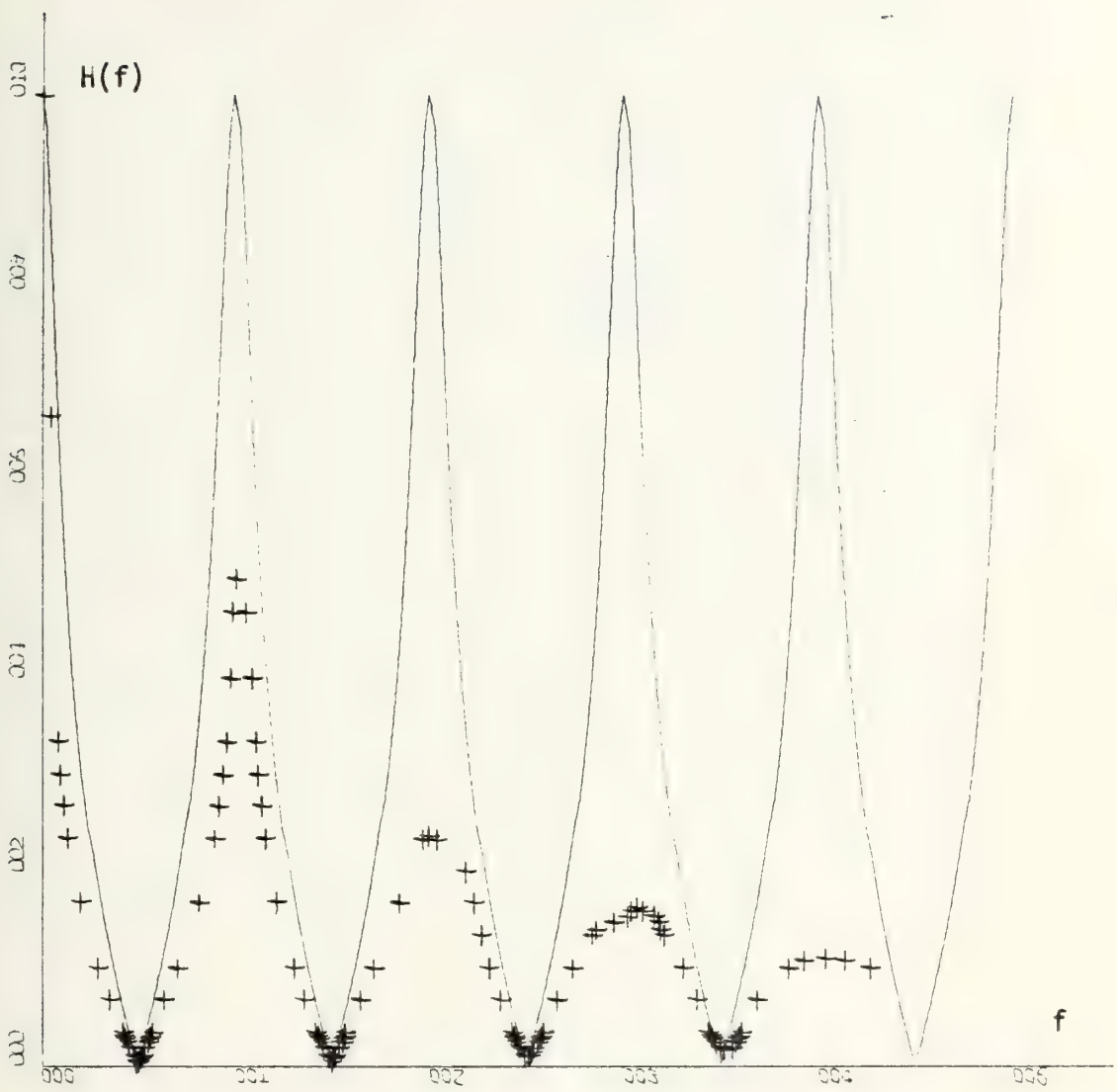


Figure 2.23. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = -0.6$ Experimental data are shown by a (+).

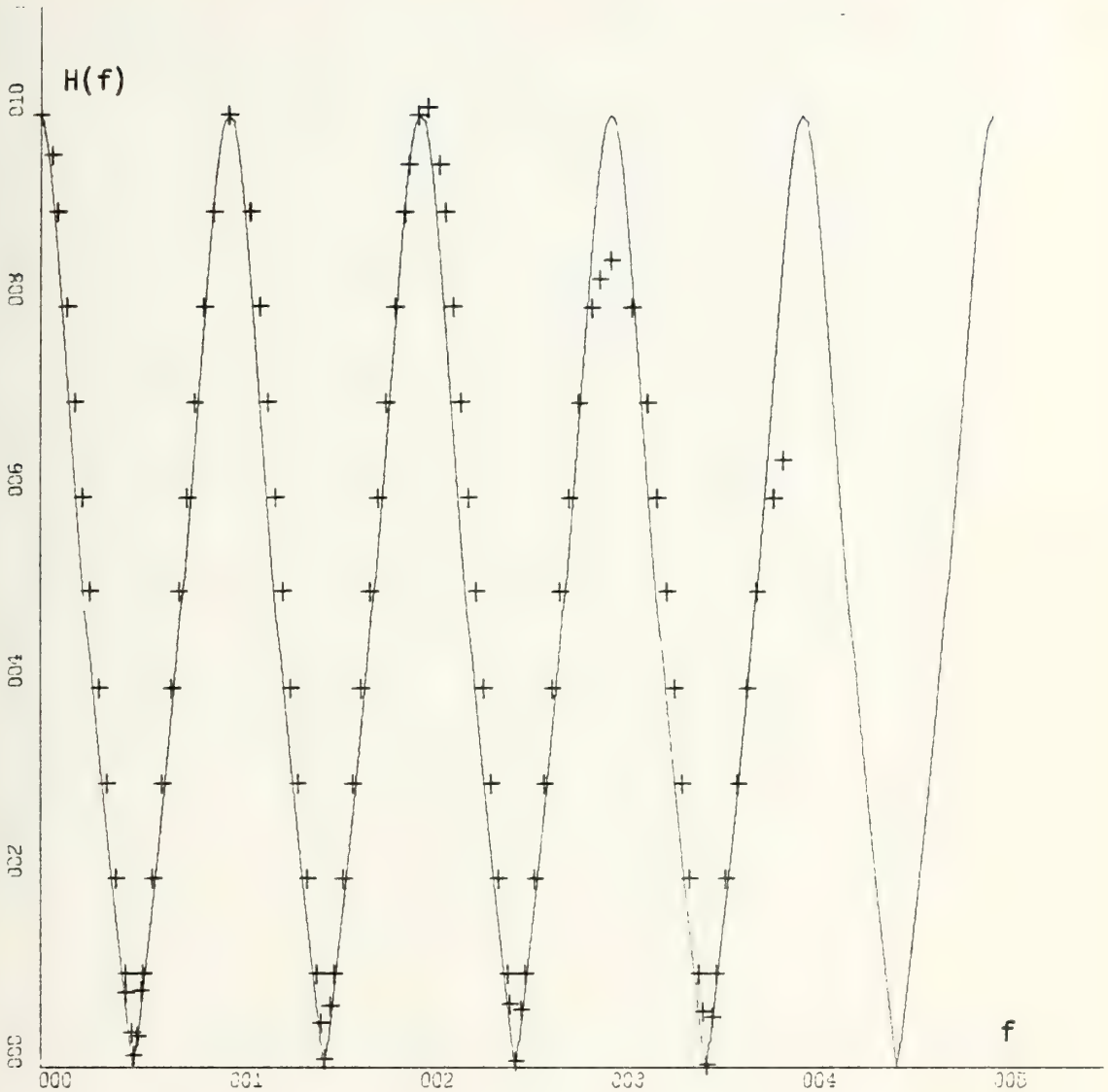


Figure 2.24. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = -0.3$ Experimental data are shown by a (+).

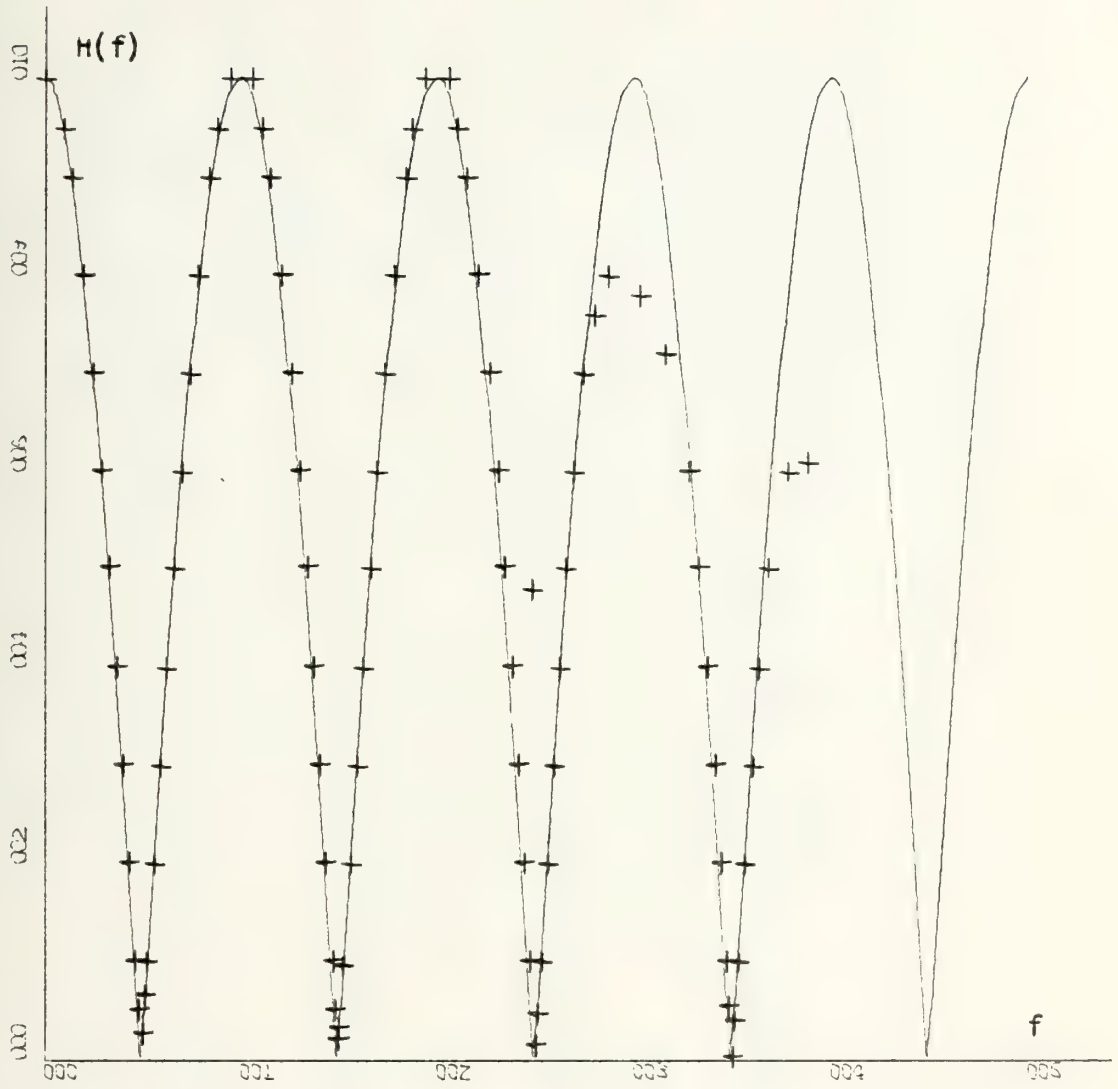


Figure 2.25. Theoretical and experimental transfer function for a first order transversal filter, $a_0 = 1$, $a_1 = 1$, $b_1 = 0$ Experimental data are shown by a (+).

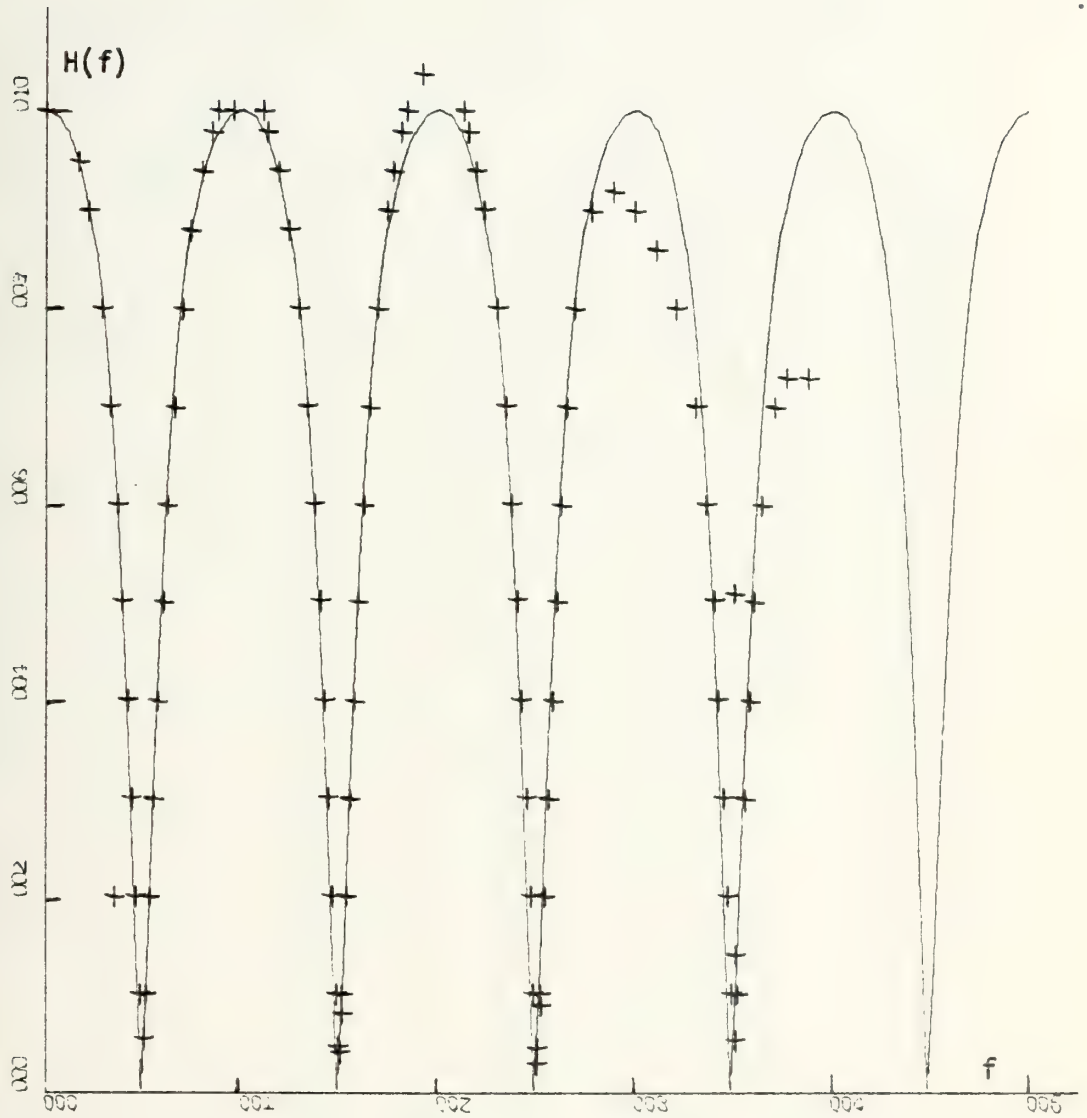


Figure 2.26. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = 0.3$ Experimental data are shown by a (+).

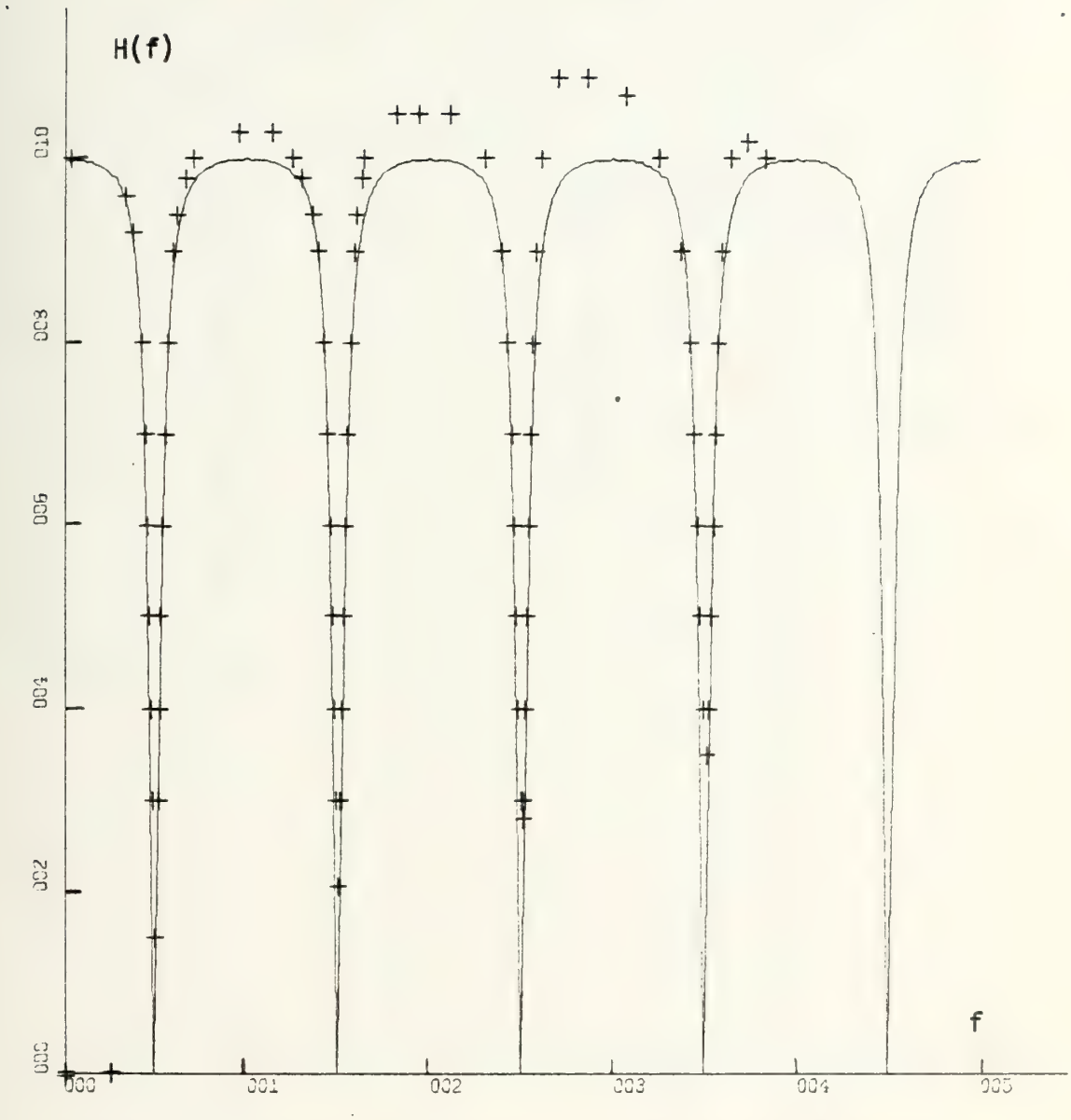


Figure 2.27. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = 0.7$ Experimental data are shown by a (+).

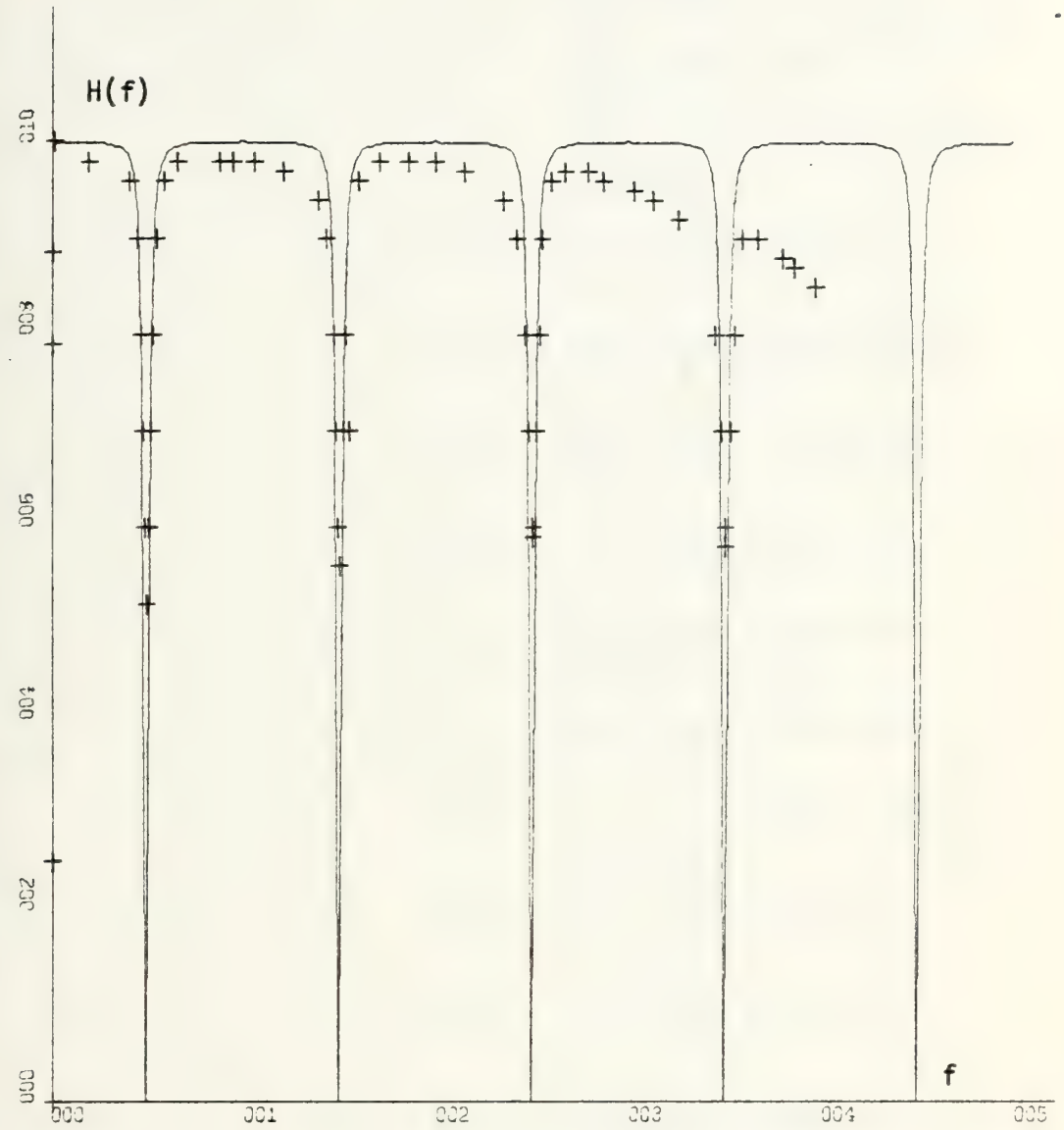


Figure 2.28. Theoretical and experimental transfer function for a first order recursive filter, $a_0 = 1$, $a_1 = 1$, $b_1 = 0.9$ Experimental data are shown by a (+).

Figure no.	A_o	A_1	B_1	Type
2.16	1	-1	0.7	High pass integrator (medium Q)
2.17	1	-1	0.3	High Pass integrator (low Q)
2.18	1	-1	0.0	High Pass (nonrecursive filter)
2.19	1	-1	-0.3	High pass canceller (low Q)
2.20	1	-1	-0.7	High pass canceller (medium Q)
2.21	1	-1	-0.9	High pass canceller (high Q)
2.22	1	1	-0.9	Low pass integrator (high Q)
2.23	1	1	-0.6	Low pass integrator (medium Q)
2.24	1	1	-0.3	Low pass integrator (low Q)
2.25	1	1	0.0	Low pass (nonrecursive filter)
2.26	1	1	0.3	Low pass canceller (low Q)
2.27	1	1	0.7	Low pass canceller (medium Q)
2.28	1	1	0.9	Low pass canceller (high Q)

On each curve, the theoretical transfer function is plotted as a continuous line. The data points marked with "+"'s are shown also. On many of the curves, agreement between theory and practice is extremely good for the first two teeth of the comb. Agreement falls off at higher

frequencies for a number of reasons and this will be taken up in the next section. However, in a few curves (no. 20, 21, and 23) the agreement does not seem to be so good. In curves 20 and 21, it appears that the coefficient b_1 was not set correctly. The correction factor mentioned in the first part of this section was used to set all the potentiometers for these measurements and it is likely that the potentiometer itself was not linear throughout its range of 20 KΩ. The data taken seems to be symmetrical about half the sampling frequencies and periodic with respect to f_s. So the error in b_1 appears to be the dominant error. In curve no. 23, we have the reverse situation ... the actual data appears to be taken for a coefficient considerably larger than 0.6. Several other curves were taken for these levels of feedback with, in general, good agreement. Figure 2.27 shows an interesting effect for which we have struggled to obtain some explanation at least qualitatively. The transfer function clearly indicates a rising amplitude response with frequency. This rising phenomenon was noticed in several curves, but not so pronounced as in no. 27. We shall consider several possible sources of this behavior in the next section.

2.2.3 Corrections in the data because of physical instrumentation

Some effects cause the data to roll off with increasing frequency. Other natural effects are periodic and cause distorted amplitude response. We shall discuss several of these

effects which are all caused by the nonideal nature of our experiments.

The most significant cause of rolloff in amplitude response of all the first order filters is the sample-and-hold effect. Nyquist's sampling theorem states that if a signal is sampled at twice its highest frequency component, it can be reproduced precisely by passage through an ideal low pass filter. We used a sample-and-hold circuit as our low pass filter. It is not ideal. The ideal filter amplitude response is shown in Figure 2.29. On the other hand, the amplitude response of the sample-and-hold circuit is shown in Figure 2.30. The general shape is sin x/ x where

$$x = \pi \, f/f_s \qquad (2\text{-}25)$$

This frequency response must be multiplied by the theoretical filter transfer function to get the projected overall transfer function corrected for sample-and-hold circuitry. This has been done for two of the curves shown in Section 2.2.2. Figures 2.31 and 2.32 show the result. As can be seen, the sample-and-hold factor reduces the theoretical calculations substantially (because most of our work is relatively close to the Nyquist frequency) and is the major reason for the frequency rolloff with increasing frequency.

The sample-and-hold factor can also be seen intuitively by considering the action in the time domain. Our clock (sampling) frequency is 20 kHz. We have taken data at 5 kHz,

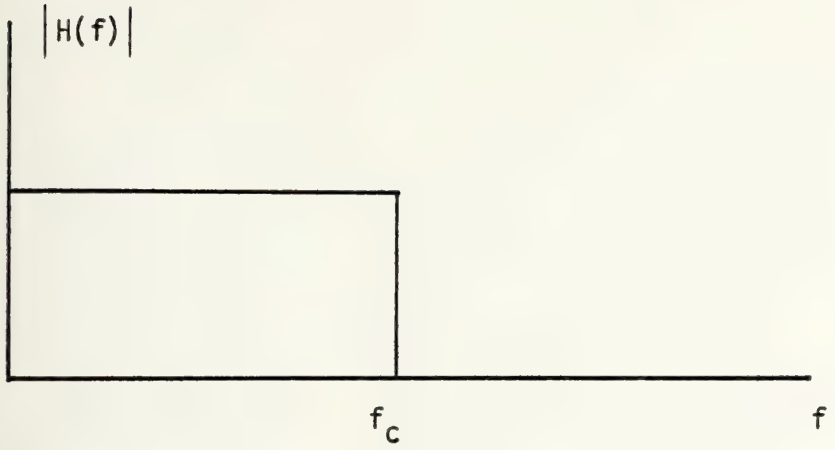


Figure 2.29. Ideal low pass filter.

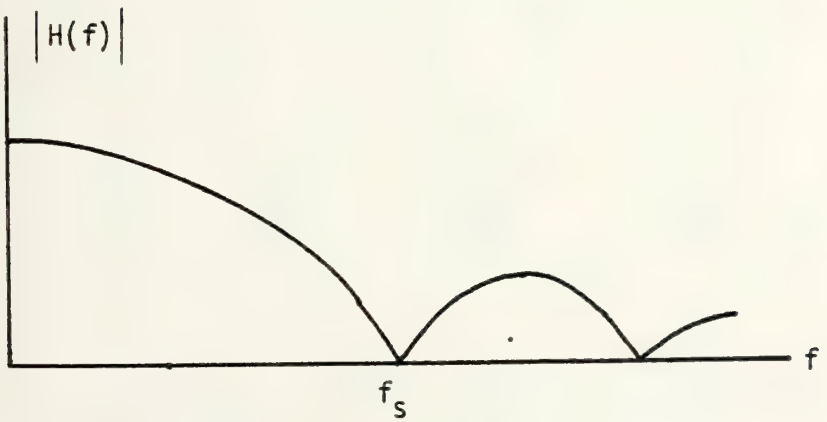


Figure 2.30. Transfer function for a sample-and-hold circuit. f_s is the sampling frequency.

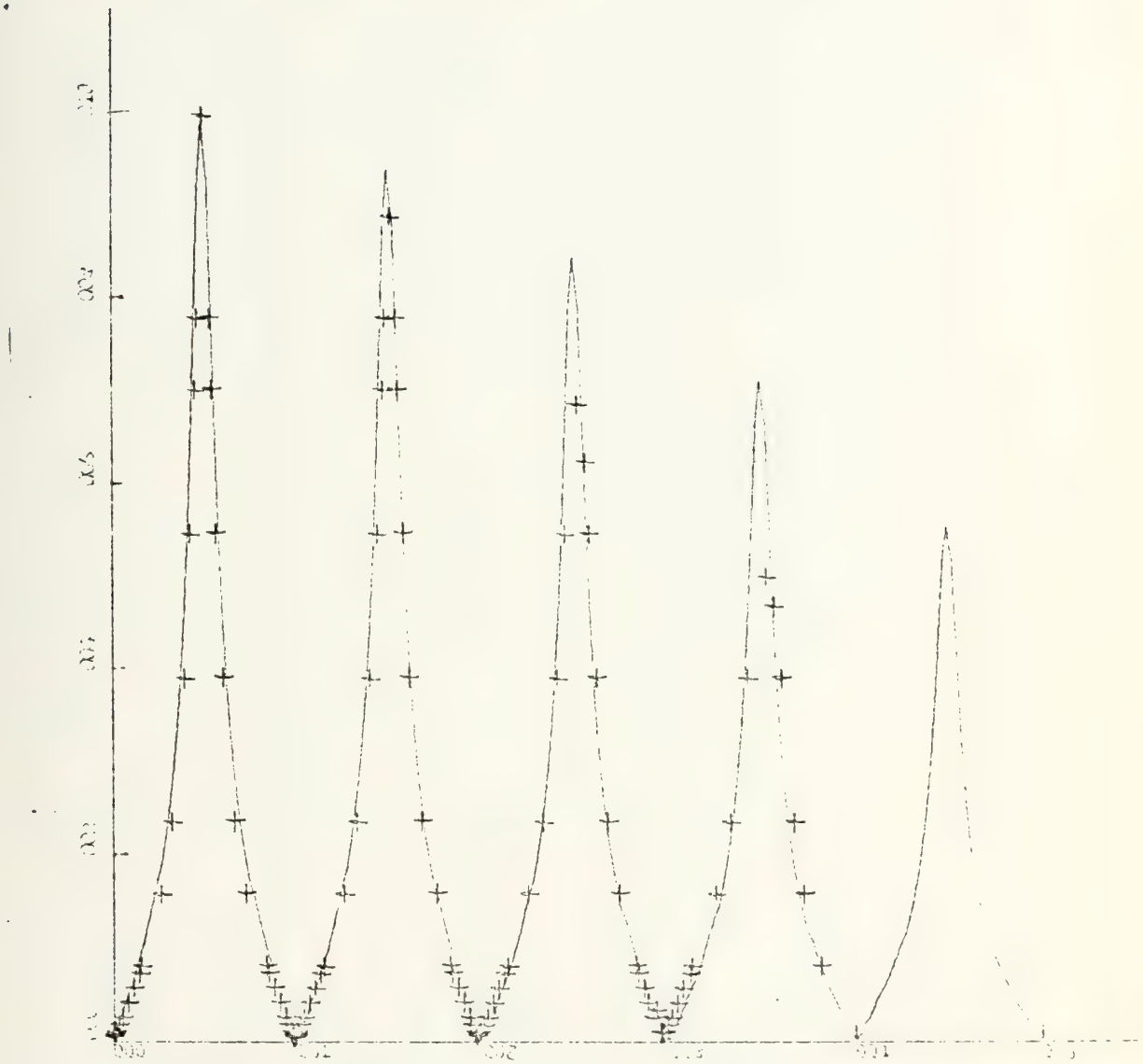


Figure 2.31 Theoretical and Experimental transfer function for a first order recursive filter, corrected for the sample-and-hold effect. $a_0 = 1$, $a_1 = -1$, $b_1 = 0.7$

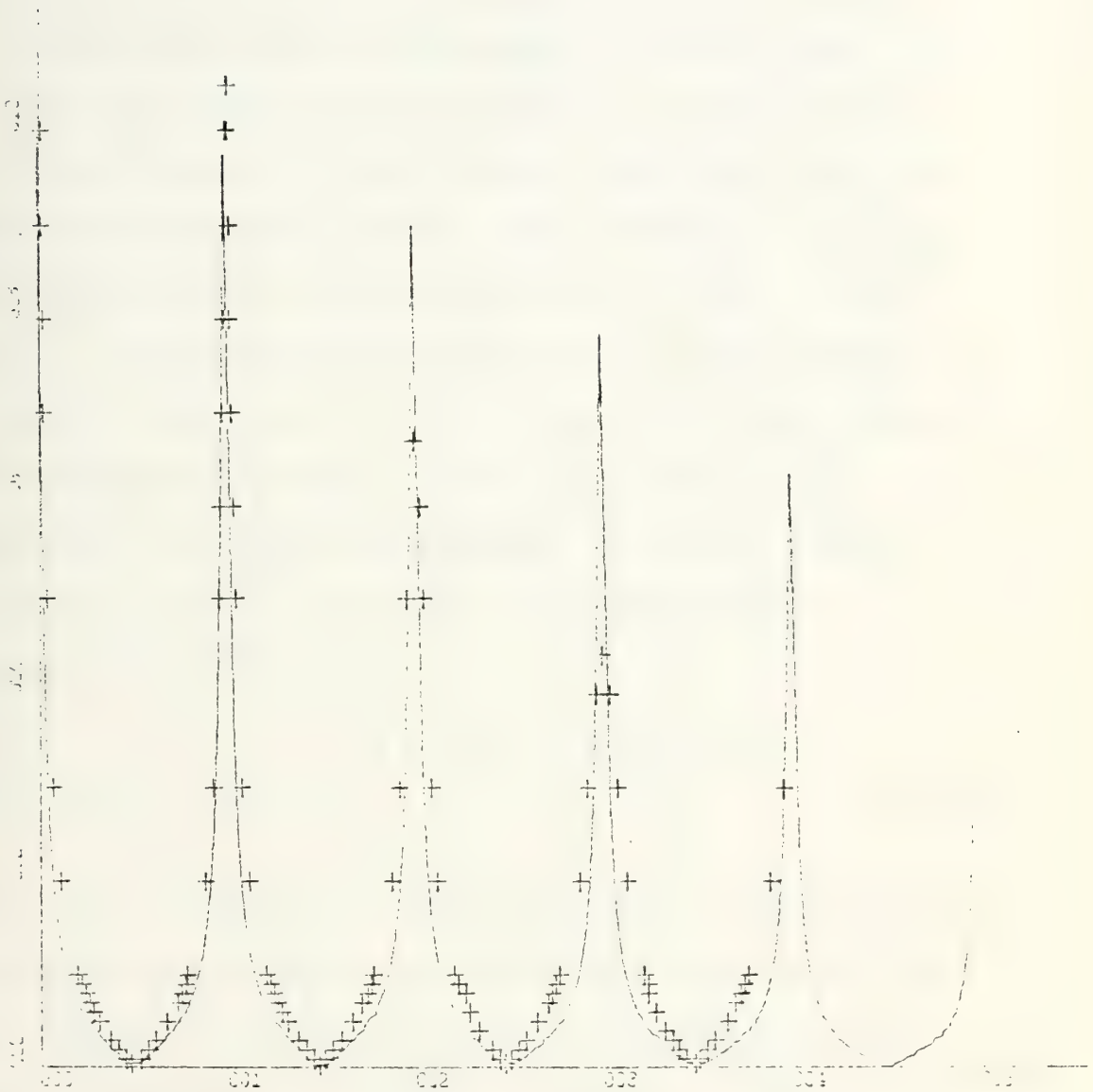


Figure 2.32. Theoretical and Experimental transfer function for a first order recursive filter, corrected for the sample-and-hold effect. $a_o = 1$, $a_1 = 1$, $b_1 = -0.9$

well below the Nyquist frequency of 10 kHz for which we ought to be able to reconstruct a signal. However, even at 5 kHz, we are taking but four samples per cycle. This effect is best seen on an oscilloscope where the four distinct levels per cycle can be observed. It is clear that the resulting waveform is not a clean 5 kHz sine wave, but instead a waveform which contains many different harmonics.

The second effect which we shall consider is the CCD efficiency. As mentioned in Chapter one, as each charge packet travels down the CCD, it is degraded slightly because of the incomplete transfer at each gate. The measure of incompleteness is called inefficiency, ε, and is commonly on the order of 10^{-4}. The factor for inefficiency is exponential and is given by

$$\text{Total inefficiency} = e^{(-n\varepsilon(1-\cos 2\pi f/f_s))} \qquad (2\text{-}26)$$

Using this expression as the attenuation of the signal after each passage through the CCD, we can calculate a new transfer function for the filter. The block diagram for explicit inclusion of this factor is shown in Figure 2.33. The overall effect of this factor is shown in Figure 2.34 for a filter with $a_o = 1$, $a_1 = -1$, and $b_1 = 0.7$ (high pass integrator). There are six curves plotted on this figure. The curve which is not attenuated at all with increasing frequency corresponds to an inefficiency of zero ... that is, the whole charge packet is transferred each time. The

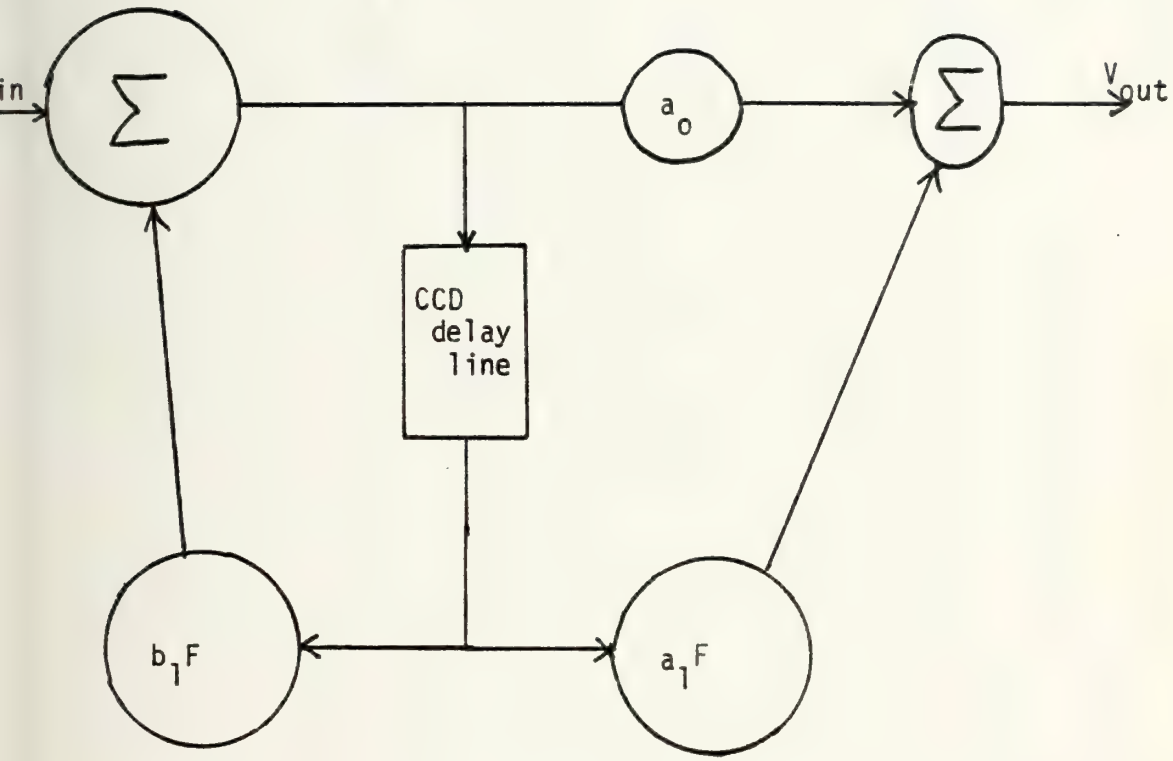


Figure 1.33. Explicit inclusion of the factor, F, for inefficiency. $F = \exp(-n\epsilon(1-\cos 2\pi f/f_s))$. First order filter.

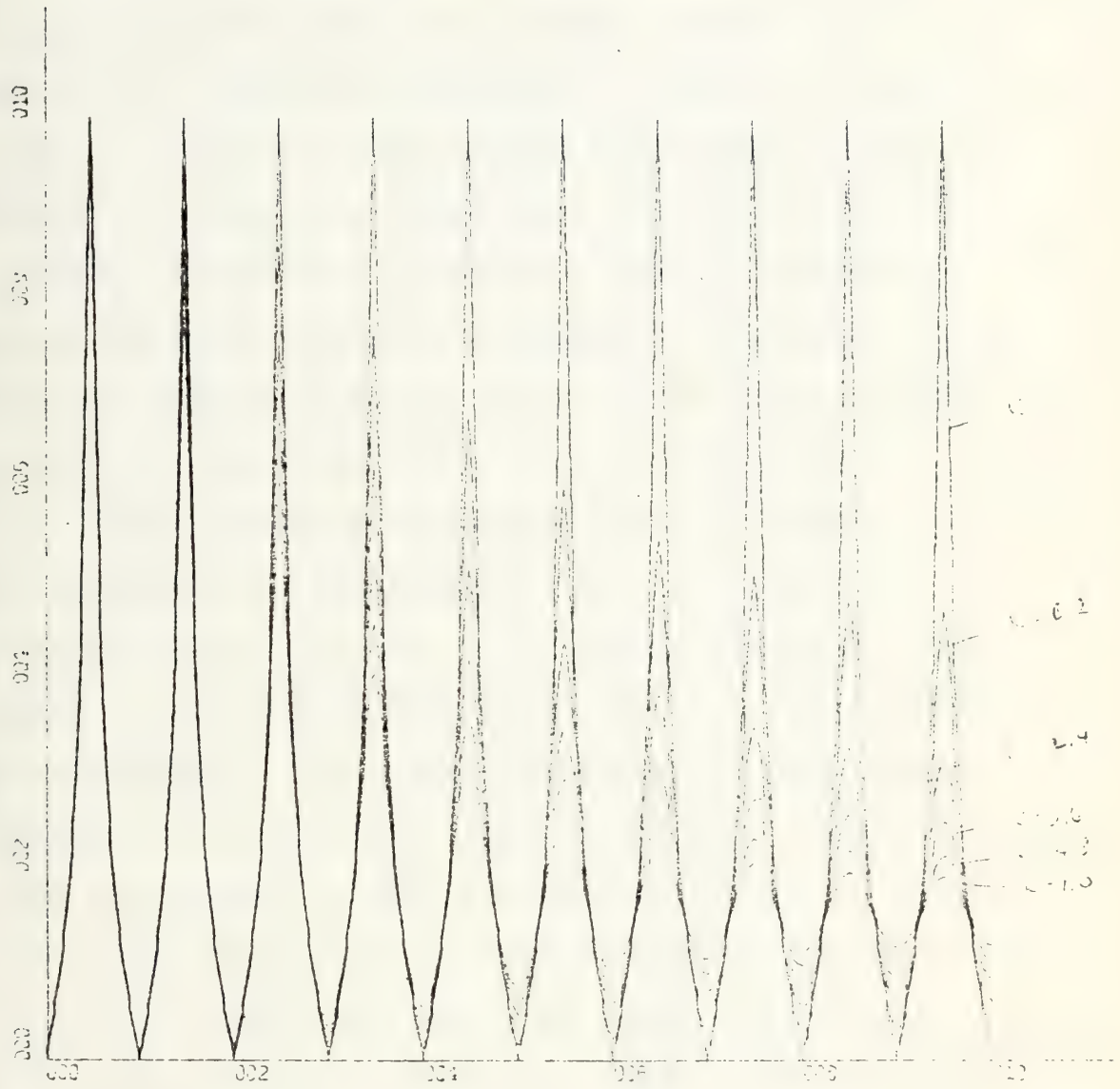


Figure 2.34. Theoretical transfer function for a first order recursive filter corrected for inefficiency. $k = n\in$, $a_0 = 1$, $a_1 = -1$, $b_1 = 0.7$

second curve down from the zero curve corresponds to a value of .0125. Each subsequent curve corresponds to an inefficiency of .0125 greater than the previous curve.

There are two important phenomena to notice in Figure 2.34. One is, of course, the falling off of the maximum as frequency is increased. The other, is just as significant, although not so easily seen, and that is the deterioration of the null (zero) as frequency is increased. This is intuitively appealing since, at low frequencies we know that $a_o = a_1$ for our filters derived by the bilinear z transform. When frequency increases to such an extent that the effect of the inefficiency has a considerable effect on the gain of the CCD, a_o no longer equals a_1. The coefficient a_1 has been shifted by the shift in gain caused by the inefficiency. If a_o does not equal a_1, then proper cancellation at the nulls does not take place. In other words, the zero which, at low frequencies, lay at $z = -1$, shifts at higher frequencies so that it lies at $z = -z_i$ where $|z_i| < 1$. Since the zero is no longer on the unit circle, we cannot expect to achieve a perfect zero in our transfer function.

Since the factor for inefficiency is periodic with respect to the sampling frequency, we expect that the shape of the curve in Figure 2.34 will be reproduced at higher regions of the spectrum. Figure 2.35 confirms this expectation. Calculations for Figure 2.35 have been made from

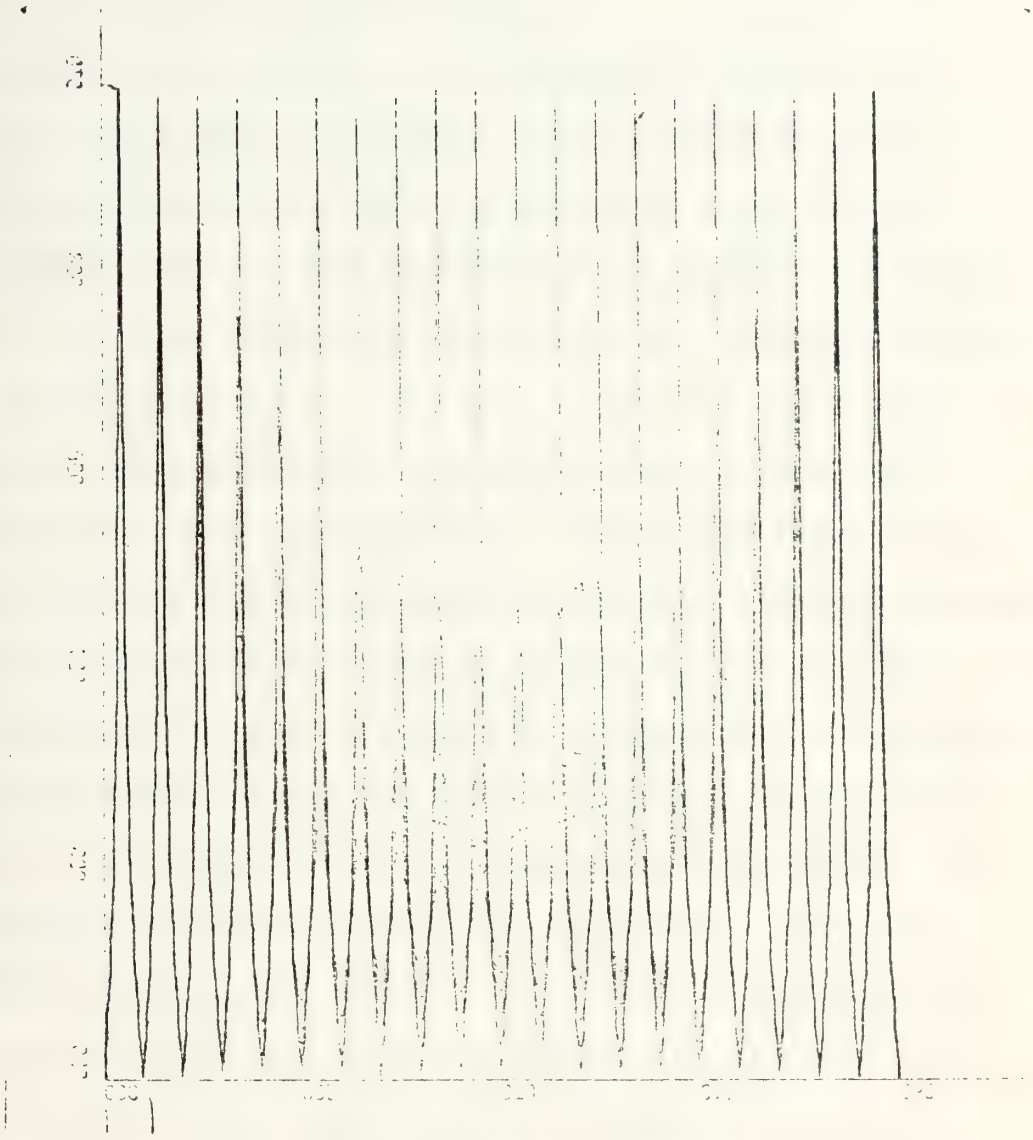


Figure 2.35 Theoretical transfer function for a first order recursive filter corrected for inefficiency, extended to f_s = 20 Khz. k = n , a_0 = 1, a_1 = -1, b_1 = 0.7

dc to f_s (20 kHz). We see the familiar symmetry about $f_s/2$, as well as the aforementioned deterioration of the zeros and attenuation of the maxima.

The obvious next step is to combine the effects of inefficiency and sample-and-hold circuitry and determine what the new theoretical transfer function ought to be. This has been done for two representative filters. Figures 2.36 and 2.37 show different aspects of the transfer function for the filter with $a_o = 1$, $a_1 = -1$, and $b_1 = 0.7$. The first figure here shows the transfer function from dc to 10 kHz corrected for both effects. The second figure shows the comparison of the experimental data with the new theoretical transfer function based on an estimated inefficiency of 10^{-3}. Agreement is good. Figure 2.38 shows the new theoretical transfer function for the filter with the coefficients $a_o = 1$, $a_1 = -1$, and $b_1 = -0.7$. Four teeth are shown. The same phenomena are evident with the additional expected nonsymmetry about $f_s/16$, $3f_s/16$, etc. This nonsymmetry is, of course, caused by the sample-and-hold and inefficiency factors, which are not periodic with respect to $f_s/8$ and are not symmetrical about $f_s/16$. The rising null can also be seen with increasing inefficiency.

One possible explanation for the rising phenomenon mentioned in the preceding section is that the potentiometers have an impedance function which falls with frequency. These potentiometers are the wire-wound type and one effect

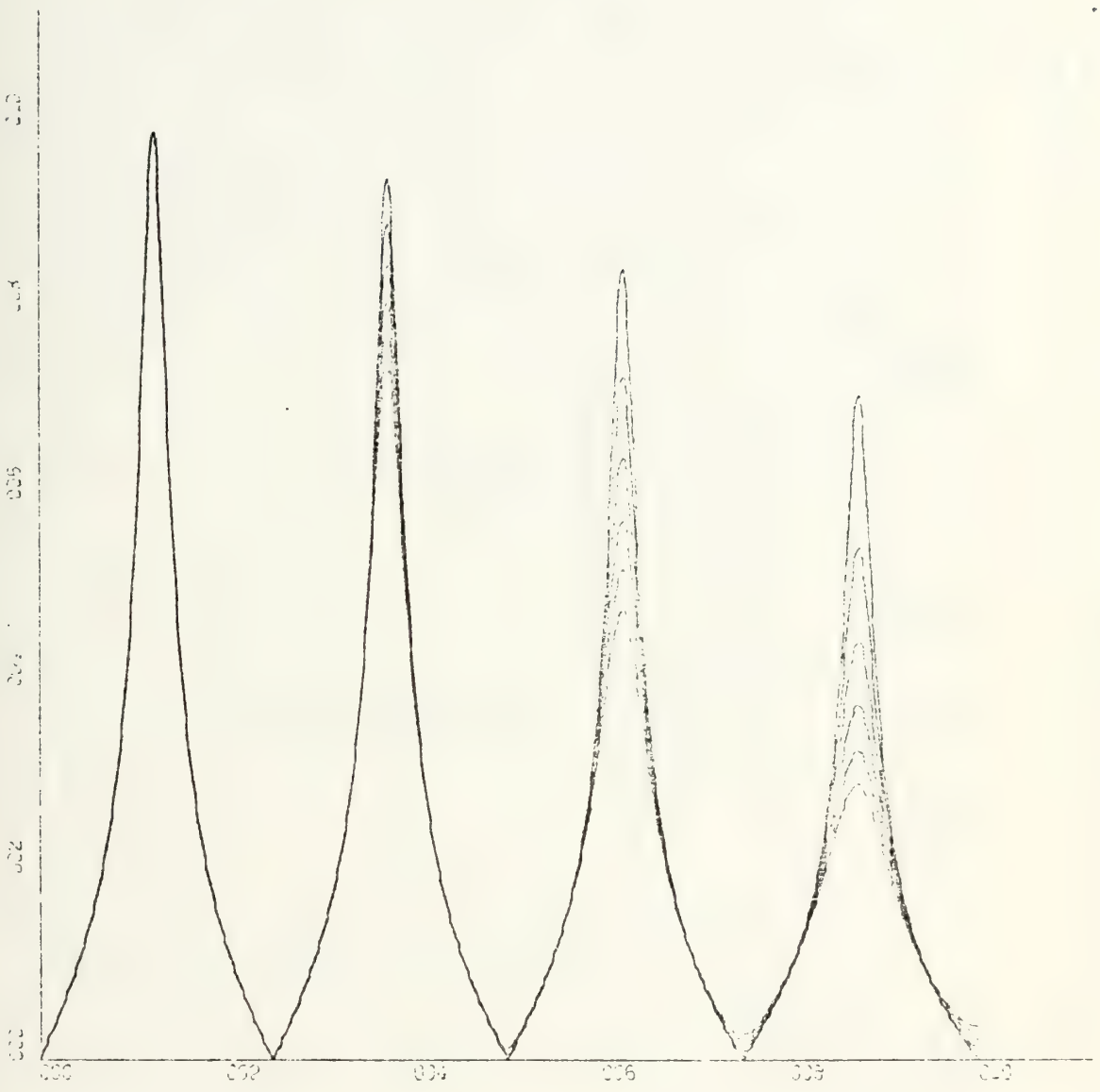


Figure 2.36 Theoretical transfer function for first order recursive filter, corrected for both sample-and-hold effects and inefficiency. Different curves are for different values of inefficiency, as explained in the text. $a_0 = 1$, $a_1 = -1$, $b_1 = 0.7$

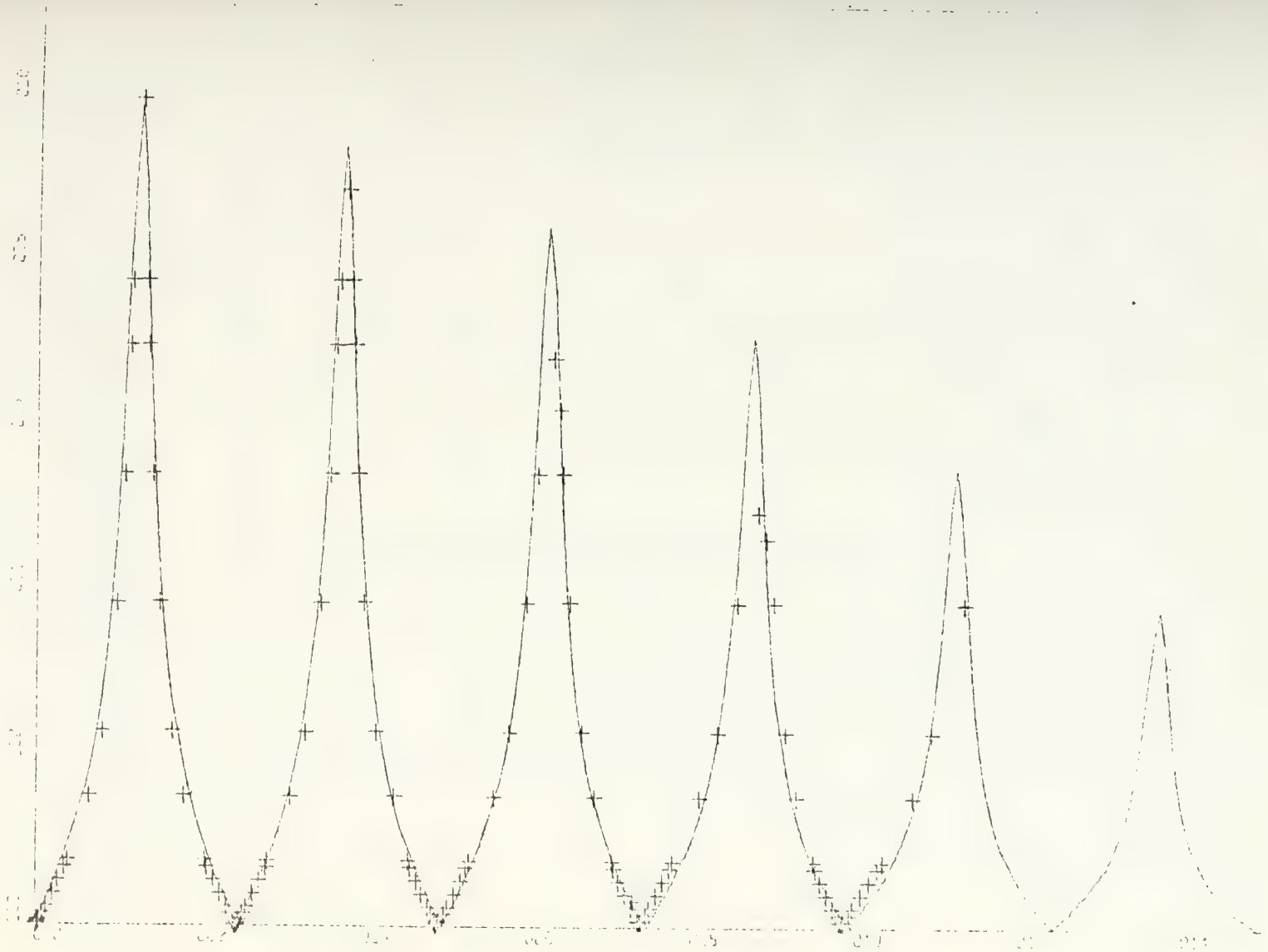


Figure 2.37 Theoretical and experimental transfer function for a first order recursive filter corrected for both sample-and-hold effects and an inefficiency of 0.001. $a_0 = 1$, $a_1 = -1$, $b_1 = 0.7$

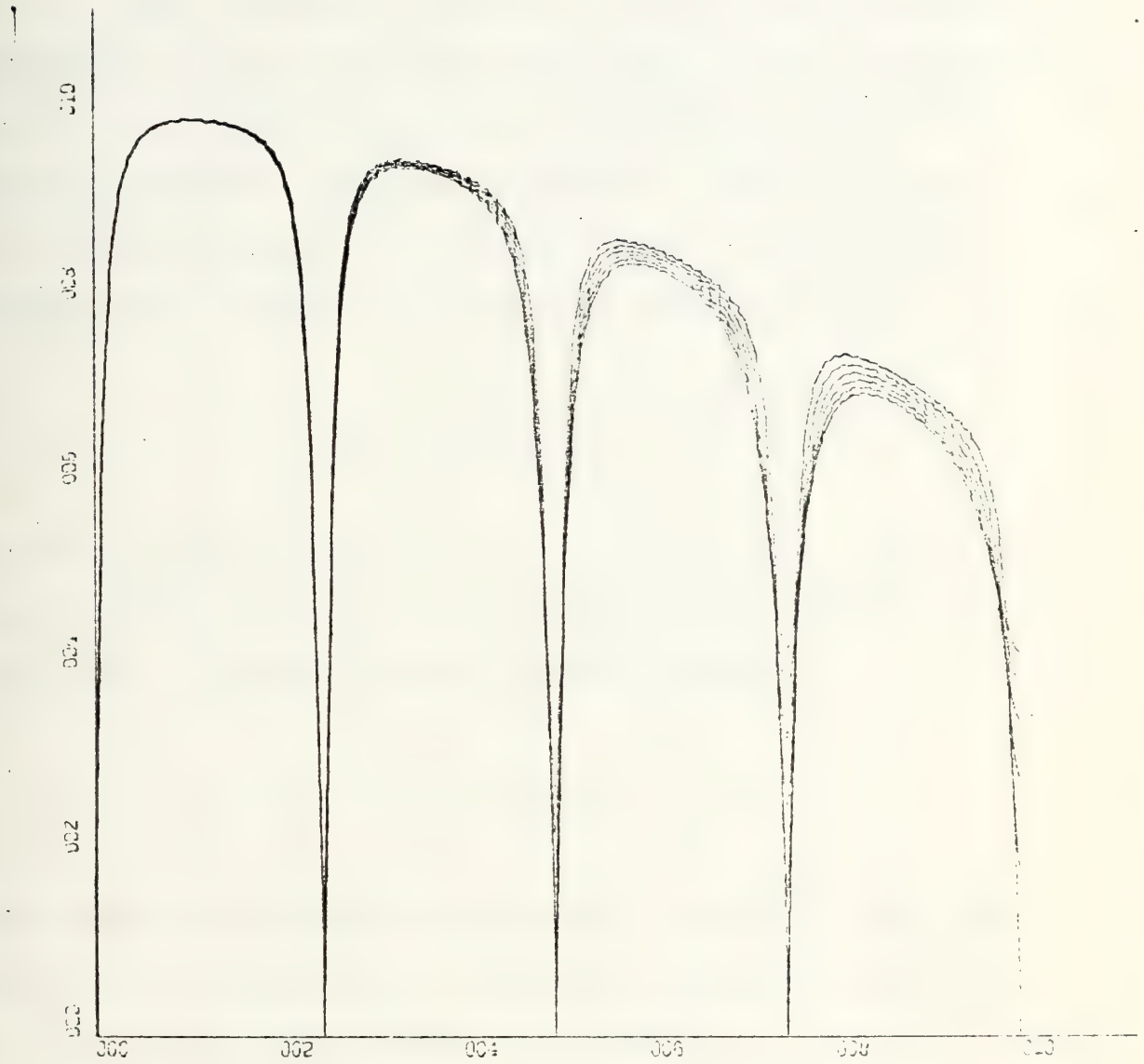


Figure 2.38 Theoretical transfer function for first order recursive filter, corrected for both sample-and-hold effects and inefficiency. Different curves are for different values of inefficiency as explained in the text. $a_0 = 1$, $a_1 = -1$, $b_1 = -0.7$

of such resistors is that they become capacitive at high frequencies due to the capacitance between adjacent loops of wire. Such a parasitic effect, along with the parasitic capacitance of the coaxial cable utilized in this experiment, would have the result of reducing the absolute value of the feedback coefficient at high frequencies. This, in turn, for settings of positive b_1 has the effect of increasing the amplitude response. For instance suppose we have

$$H(z) = \frac{1 + z^{-1}}{1 + 0.3\ z^{-1}}$$

The absolute value of H(z) at dc is 2/1.3 or 1.54. At higher frequencies, when 0.3 is reduced by a factor of 0.7, we have the following transfer function value:

$$H(z) = \frac{1 + z^{-1}}{1 + 0.21\ z^{-1}} = 1.65$$

which represents an increase of about 7% which is about what is observed in Figure 2.26. For higher values of feedback, say 0.7, we have the following two values:

$$H(0) = 1.18 \qquad H(f_x) = 1.34$$

where the increase is about 14% which corresponds well with the actual observed data in Figure 2.27.

Strangely, the phenomenon did not occur for $b_1 = 0.9$. It is possible that the coaxial connections were changed for the data obtained at this high feedback setting.

To summarize this section, we have discussed the following factors which cause deviations of experiments from theory:

A. Sample-and-hold --- causes attenuation of maxima and nonsymmetry of comb filter

B. Inefficiency --- causes attenuation of maxima and deterioration of zero (null). Also causes nonsymmetry of comb, but its effect on symmetry is not so marked as the sample-and-hold effect.

C. Potentiometer --- causes an increasing of amplitude response for positive feedback cases with increasing frequency. Causes a decreasing of amplitude response for negative feedback case with increasing frequency.

2.2.4 Selectivity

The idea of selectivity is central to filtering. A filter changes the frequency content of a signal. Highly selective filters eliminate almost all frequency components except the small band which they pass. No discussion of filters is complete without some mention of selectivity.

The basic measure of selectivity is Q. Q is defined in various places by various equations, but it is appropriate here to use the definition

$$Q = \frac{f_{center}}{bandwidth} \qquad (2\text{-}27)$$

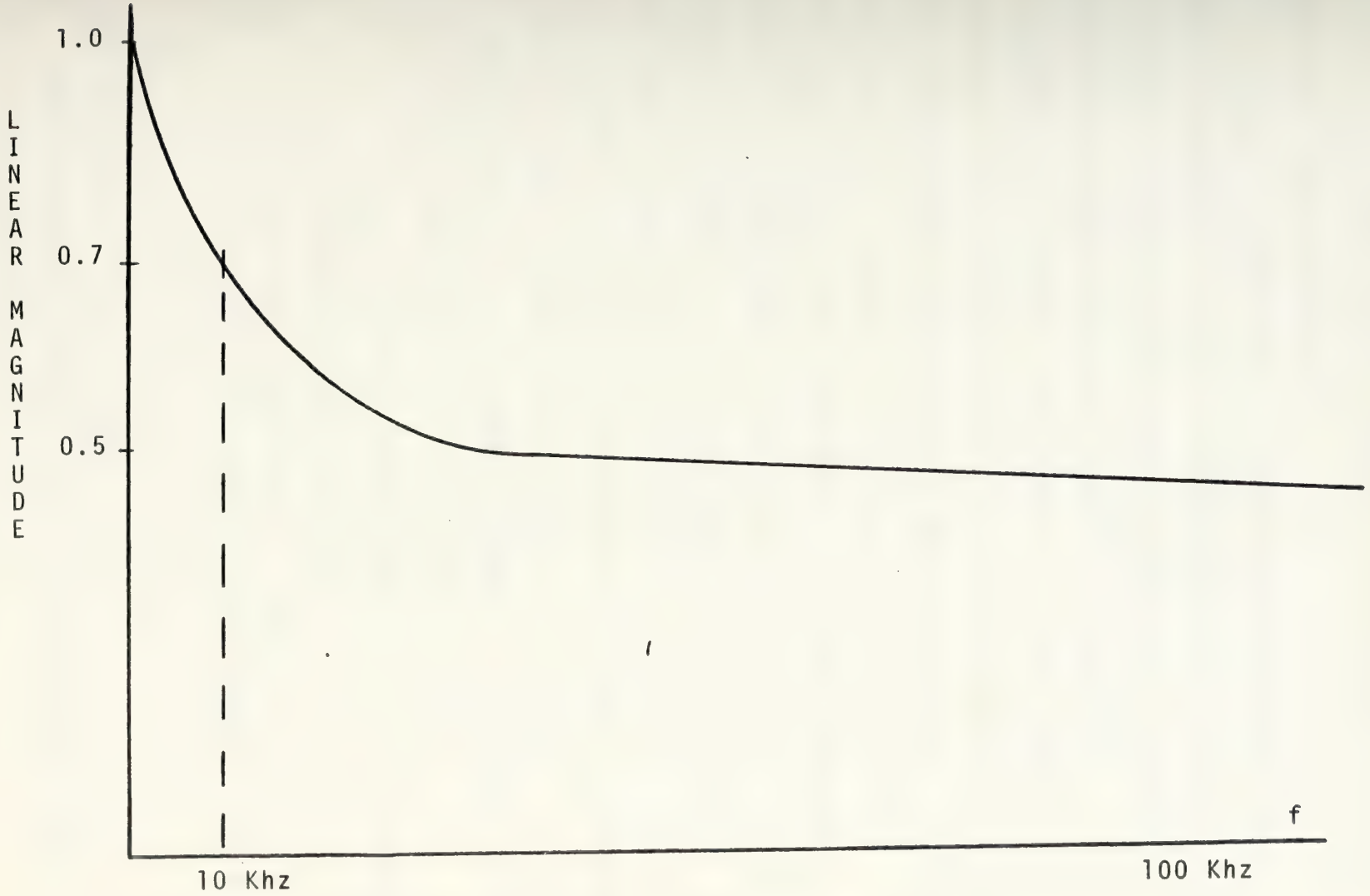


Figure 2.39. Frequency response of potentiometers.

where f_{center} refers to the center frequency of the passband and the bandwidth is the distance (in frequency) between the two 3 db down points.

The above definition is fine for integrators. The center frequency is well defined and the bandwidth is small (usually) relative to the center frequency. The above definition breaks down when applied to cancellers. In canceller type filters, the interest is not really in the passband, but in the stopband. A known noise signal exists, is received, and filtered. The canceller ideally stops all the noise. Of course, in practice, this is not possible for some of the noise components are in the passband of the filter. Nevertheless, if we have a good idea of the spectrum of the noise, then we can build a filter which is optimum for the noise under consideration. At first, we attempted to use the reciprocal of the definition given in Equation (2-27). This did not work so well since it made Q a nonlinear function of the center frequency. A better choice of Q to be used with cancellers appears below:

A. H(f) is the transfer function of a canceller
B. Normalize H(f) to unity
C. Form the difference 1 - H(f) = H'(f)
D. Now utilize Equation (2-27) and find the Q of H'(f).

This is the Q of H(f) as well.

This procedure retains the linearity and has worked well in our calculations.

Figure 2.40 shows the theoretical calculation of Q for a high pass filter derived by bilinear z transform with variable feedback coefficient. As can be seen, the Q is low for negative values of b_1 and also for positive values of b_1 less than about 0.7. Q rises dramatically for high values of b_1 (near 1). Of course, the usual definition of Q was used in this calculation. The use of the 1 - H(f) definition would produce the identical curve except that the rise in Q would occur near -1. The overall evaluation of a design should use the appropriate definition of Q.

2.2.5 Summary

In this section, we have presented our experimental results together with theoretical calculations and some reasons why the experimental data do not always agree exactly with the theory. The first order filter has the advantage of simplicity and the use of the bilinear z transform often provides us with a rather large maximum-to-minimum response ratio. One of our conclusions is especially significant: The standard z transform is a rather poor tool to use when designing filters or other processors with a frequency response specification in mind. Since most design is done to frequency specifications, we would strongly suggest that the standard z transform not be chosen for a means to design unless the sampling frequency is much much greater than the sampled analog cutoff frequency desired. Several factors interfere with perfect operation of our filters:

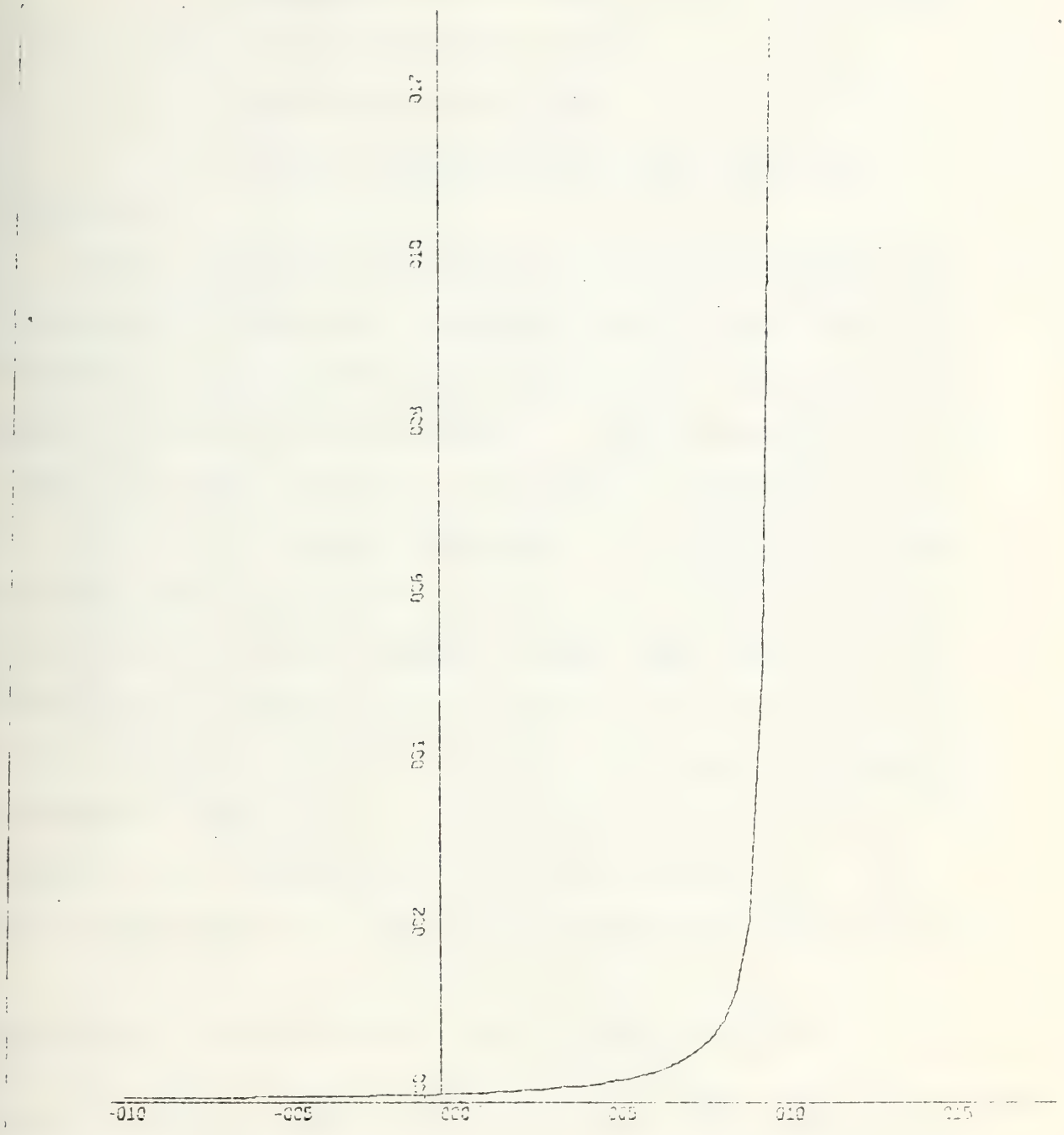


Figure 2.40 Quality factor (Q) as a function of feedback coefficient for a high pass first order recursive filter. $a_0 = 1$, $a_1 = -1$

A. Instrumentation problems such as the frequency response variation of the potentiometers.

B. Inefficiency of the CCD.

C. Nonideal response of the sample-and-hold circuitry.

Figure 2.41 and 2.42 are of interest. They summarize the theory of first order sampled analog filters designed with the bilinear z transform. Both figures show the same thing, but from different aspects. The figures show the wide range of transfer functions available to a designer merely by changing the feedback coefficient, b_1, of a first order filter. The x-axis is the b_1 value, the y-axis is frequency, and the z-axis is, of course, the amplitude response. The coefficient b_1 varies from -0.9 to 0.9 while frequency varies from dc to $2.5f_s$. Where the curves have especially wide passband is where $b_1 = +0.9$ (foreground in Figure 2.41, background in Figure 2.42). The other extreme ($b_1 = -0.9$) is characterized by a very narrow passband and wide stopband.

SECTION 2.3 THE RECURSIVE SECOND ORDER COMB FILTER

2.3.1 The basic characteristics of the second order comb filter are similar to those of the first order except that two poles and as many as two zeros may occur in the transfer function. The sampled analog transfer function is

$$H(z) = \frac{a_o + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}} \qquad (2\text{-}28)$$

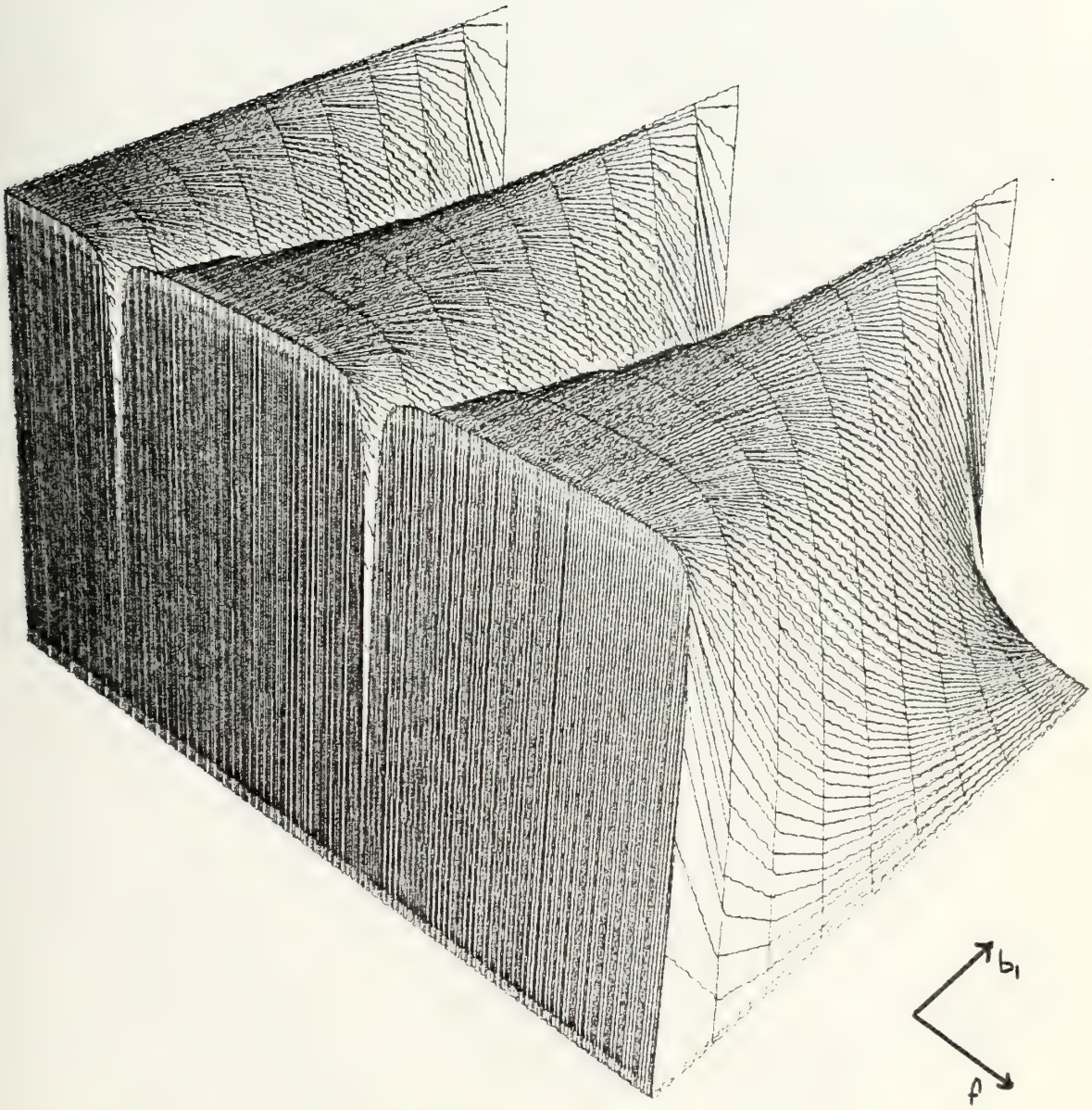


Figure 2.41. Low pass theoretical transfer function for a first order recursive filter with various values of feedback, $a_0 = 1$, $a_1 = 1$

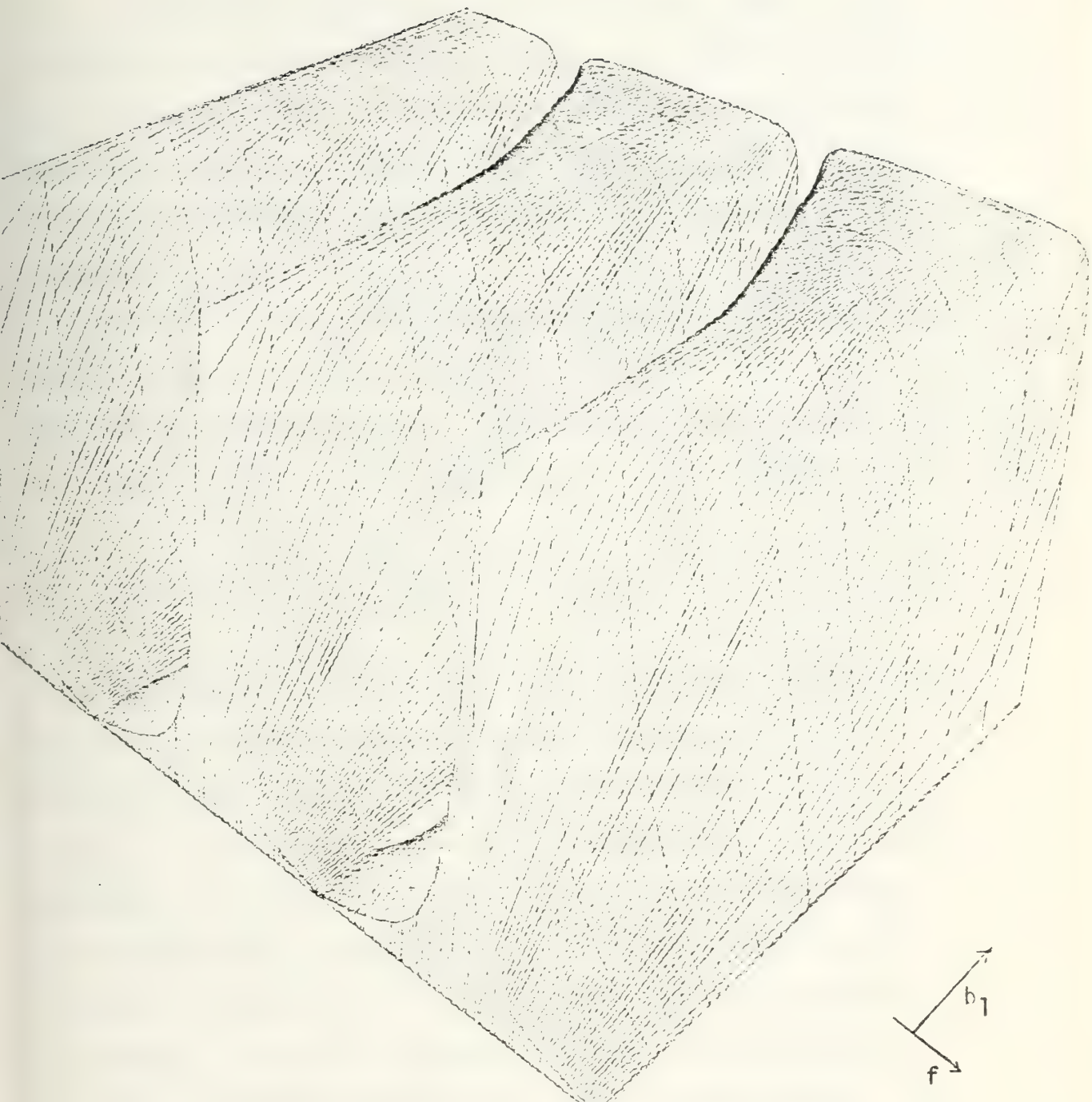


Figure 2.42 Low Pass theoretical transfer functions for a first order recursive filter with various values of feedback. $a_0 = 1$ $a_1 = 1$

The second order recursive filter is widely used for such applications as radar signal processing, bandpass and band-stop filtering, and in communications. It is more widely used than the first order filter since it offers a wider variety of transfer function shapes for a given sampling frequency.

The comb filter design technique, as in the first order case, utilizes a serial-in serial-out delay line, such as the CCD discussed earlier. The true transfer function for our experiments using two eight stage delay devices is

$$H(z) = \frac{a_o + a_8 z^{-8} + a_{16} z^{-16}}{1 + b_8 z^{-8} + b_{16} z^{-16}} \qquad (2\text{-}29)$$

Such a transfer function is clearly a sixteenth order filter, one with, at most, sixteen poles and sixteen zeros. As before, there are many coefficients missing: $a_1, a_2, \ldots, a_7, a_9, \ldots, a_{15}$ and the same corresponding b's. This design procedure enables the engineer to consider his sixteenth order filter as a second order filter where his adjusted sampling frequency is one eighth that of the true sampling frequency. As shown in Appendix C, the poles and zeros for such a transfer function are not equidistant as in the first order case. Figure 2.43 shows that the poles and zeros are arranged so that they are symmetrical with respect to dc, $f_s/8$, $2f_s/8$, ..., $nf_s/8$. Such a symmetric nature yields a transfer function periodic with respect to $f_s/8$.

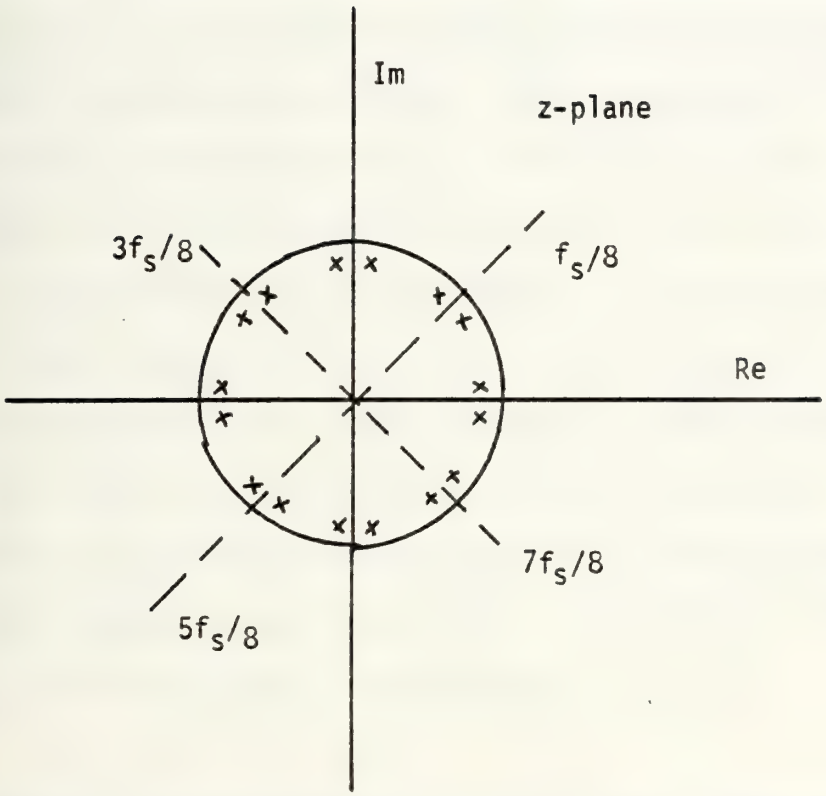


Figure 2.43. Symmetrical pole plot on the z-plane. Poles are spaced symmetrically to radial lines of the form $nf_s/8$, where $n=0,1,2,\ldots,7$. This affords the basis by which we make the sixteen-to-second order translation.

The ability to regard a sixteenth order filter as a second order filter is important and significant for the same reasons which applied to the eighth-to-first order translation. However, a second order filter is not so simple as a first order filter. The effect on the transfer function of a variation of one or more coefficients is not generally intuitively obvious. For instance, an engineer may need to know the effect of the variation of b_1 and b_2 by some small percentage because of thermal variations. Such parameters as cutoff frequency, "skirt slope", and ripple must remain within system tolerances, even though the coefficients vary. An analysis of these effects has been made and will be explained fully in Section 2.4. The analysis is for second order recursive sampled analog filters. The same analysis undertaken for a sixteenth order filter would, of course, be a huge project.

Section 2.4 DESIGN METHODS FOR A SECOND ORDER SAMPLED ANALOG FILTER

2.4.1 Introduction

The design of second order filters is necessarily more complex than the design of first order filters. The first order filter had but two possible singularities, one pole and one zero, and, moreover, each was by necessity, real. Thus, we were dealing with singularities restricted to the real line on the z-plane, and design was essentially a

one-dimensional problem ... deciding where on the real line to place the singularities. Obvious rules were available ... as shown in the experimental and theoretical figures of Section 2.2, the effects on the shape of the transfer function by the position of the pole and zero become intuitive and easy to work with for the design engineer. However, the relationship between the shape of the transfer function and the position of the poles and zeros of a second order filter is not so well understood. Of course, as the order of the filter increases, the effect of the placement of the poles and zeros becomes more and more complex.

Therefore, the design procedure for second order and above can no longer rely on intuitive means. Analytical processes are required for the proper design of second order filters. There are to be discussed here three different methods of designing such filters:

A. Replication in the z-domain of an analog filter whose characteristics in the s-domain are well known and well documented.

B. Selection of coefficients by reference to charts which indicate the positions of the relevant parameters: cutoff frequency, ripple values, etc. In this section we will derive and discuss the b-plane triangle of stability for second order sampled analog filters.

C. Direct sampled analog filter design using cost functions and the general purpose digital computer (Fletcher-Powell algorithm).

Section 2.4.2 Replication in the z-domain of known analog filters

Three of the best known analog filters are the following: Butterworth, Chebyshev, and elliptical (sometimes called Cauer parameter filters). Briefly, the Butterworth filter has a frequency characteristic which is monotonic in both passband and stopband, has steeper skirts for higher orders, and has a maximum error (deviation from ideal) in each stopband and passband which occurs at one and only one frequency. This maximum error gave rise to the thought that it might be possible to "spread" this error throughout a range of frequencies, and thus make the error at any one frequency much less. This was accomplished by the use of polynomials in the denominator of the transfer function which were called Chebyshev polynomials. Polynomials which "spread" the error throughout the passband only have roots (poles of the transfer function) which lie on an ellipse in the s-plane. This location of poles results in a designed-in ripple effect in the passband. The size of the ripple is one of the parameters of the filter and is held constant by a specification, usually called ε. This specification often determines the order of the filter required since higher orders can reduce the size of the ripple.

The elliptic filter takes advantage of these polynomials to produce a ripple in the stopband as well. In other words, it spreads the overall error in the stopband over a wide range of frequencies.

2.4.2.1 The Butterworth Filter

The Butterworth filter is characterized by having a monotonic transfer function in both the passband and stopband. The squared magnitude transfer function is given by

$$|F(\omega)|^2 = \frac{1}{1 + (\frac{\omega}{\omega_c})^{2n}} \qquad (2\text{-}30)$$

where n is the order of the filter. The design procedure is straight-forward. The poles are located on a circle of radius ω_c and are equally spaced. If the order is odd, then one pole lies on the negative real axis. If the order is even, then the negative real axis lies midway (angularly) between two poles.

For first order, there is one pole and no zeros. The transfer function was considered in Section 2.1. Both the standard z transform and the bilinear z transform were applied to it to derive a replication in the z-domain. As was noted at the time, the standard z transform produced impulse invariance while the bilinear z transform insured that the ranges of the two transfer functions (the analog and the sampled-analog) would be precisely equal over the domain of real frequencies. Warning was given for the standard z transform because a designer could not be sure at the outset whether the cutoff frequency, known for the analog filter, would even exist in the sampled analog filter. Expressions for the cutoff frequency were given in terms of the coefficient, b_1, and inversely.

The second order Butterworth filter can be transformed in exactly the same ways ... either by standard z transform or the bilinear z transform. The second order Butterworth polynomial has a transfer function of the form

$$H(s) = \frac{1}{s^2 + \sqrt{2}\, s\omega_c + \omega_c^2} \qquad (2\text{-}31)$$

The poles of the transfer function lie at the following positions:

$$s = -\frac{\omega_c}{\sqrt{2}} \pm j\,\frac{\omega_c}{\sqrt{2}} \qquad (2\text{-}32)$$

Once these things are known, the design of the corresponding sampled analog filter can proceed. First we consider design by the impulse invariant method. The first step is to find the residues and form

$$H(s) = \sum_{i=1}^{2} \frac{A_i}{s + s_i} \qquad (2\text{-}33)$$

Then the inverse LaPlace transform is found as before

$$h(t) = \sum_{i=1}^{2} A_i\, e^{-s_i t} \qquad (2\text{-}34)$$

This expression is sampled at some interval, T, and the Z transform taken of the resulting sampled signal:

$$Z(h'(n)) = \frac{2}{\beta} \frac{e^{-\alpha T} \sin \beta T \, z^{-1}}{1 - e^{-\alpha T} 2\cos \beta T \, z^{-1} + e^{-2\alpha T} z^{-2}} \qquad (2\text{-}35)$$

which means that the coefficients of the canonical form for a sampled analog filter are

$$a_o = 0 \qquad b_o = 1$$

$$a_1 = \frac{2}{\beta} e^{-\alpha T} \sin \beta T \qquad b_1 = -2e^{-\alpha T} \cos \beta T$$

$$a_2 = 0 \qquad b_2 = e^{-2\alpha T}$$

In order to find the cutoff frequency of the analog transfer function, we form the equation

$$|H(\omega_c)|^2 = |H(0)|^2/2 \qquad (2\text{-}36)$$

These equations must be solved. The results are given in Equation (2-37):

$$f_c^2 = \frac{1}{2\pi} \omega_c \left(-\frac{1}{2} \pm \frac{1}{\sqrt{2}}\right) \qquad (2\text{-}37)$$

where the factor ω_c is the same quantity as in the Butterworth polynomial. The roots of this expression are four in number including one positive real root, one negative real, one positive imaginary, and one negative imaginary. We naturally take the positive real root as our cutoff frequency.

For the sampled analog case, the cutoff frequency cannot be calculated so easily. The form for the equation relating

the cutoff frequency to the coefficients can be derived and solved but it is tedious. The expression is a quadratic and includes the eigenvalues of the analog transfer function as parameters:

$$f_c = \frac{1}{2\pi T} \cos^{-1}\{ \frac{\cos \beta T + \cos \beta T e^{-2\alpha T} \pm \sqrt{\cos^2 \beta T(1+2e^{-4\alpha T}+e^{-6\alpha T})}}{2e^{-2\alpha T}}$$

$$\cdot \frac{+ \cos \beta T[-8e^{-\alpha T} + 4e^{-2\alpha T} -8e^{-3\alpha T}]+1+10e^{-2\alpha T}+e^{-4\alpha T}}{2e^{-2\alpha T}}\}$$

(2-38)

This expression, though complicated, is straightforward. Simply insert the values of α and β which are the coordinates of the poles and zeros in the s-plane and compute. That these values are always known is inherent in the ability to design an analog Butterworth filter. Naturally, the solution to this equation can perhaps be done best by a computer. A program for its solution was written and tested.

In order to use the bilinear transformation, we need not find the residues or find the inverse LaPlace transform of the analog transfer function. We make the substitution as we did in Section 2.1 for a first order filter:

$$s = \frac{2}{T} \frac{z - 1}{z + 1}.$$

The second order Butterworth filter, transformed via the bilinear z transform into the z-domain produces the following results:

$$H(z) = \frac{G(z^{-2} + 2z^{-1} + 1)}{1 + \left(\frac{-4+2\alpha^2+2\beta^2}{4+4\alpha+\alpha^2+\beta^2}\right) z^{-1} + \left(\frac{4-4\alpha+\alpha^2+\beta^2}{4+4\alpha+\alpha^2+\beta^2}\right) z^{-2}} \qquad (2\text{-}39)$$

where G is the overall gain of the filter and is given by

$$G = \frac{T^2}{4 + 4\alpha + \alpha^2 + \beta^2} \qquad (2\text{-}40)$$

The coefficients of the sampled analog filter are easily picked off the transfer function:

$$a_o = 1 \qquad b_o = 1$$

$$a_1 = 2 \qquad b_1 = \frac{-4+2\alpha^2+2\beta^2}{4+4\alpha+\alpha^2+\beta^2}$$

$$a_2 = 1 \qquad b_2 = \frac{4-4\alpha+\alpha^2+\beta^2}{4+4\alpha+\alpha^2+\beta^2}$$

where α and β are the real part and the imaginary part of the pole pair (s-plane) respectively.

The cutoff frequency can be found in the usual manner. Like the result in the impulse invariant approach, the cutoff frequency calculation turns out to be complicated. It is a function of the singularities in the s-domain. If we let the cutoff frequency be a function of the sampled analog coefficients then we obtain a quadratic equation in $\cos(2\pi f_c/f_s)$ for which the following are coefficients:

term	coefficient
$\cos^2(2\pi\ f_c/f_s)$	$1 + b_1^2 + b_2^2 + 2b_1b_2 + 2b_1 - 6b_2$
$\cos\ (2\pi\ f_c/f_s)$	$2 + 2b_1^2 + 2b_2^2 + 4b_2$
$\cos^0(2\pi\ f_c/f_s)$	$-1 - b_1^2 - b_2^2 + 2b_1b_2 + 2b_1 + 2b_2$

(2-41)

Notice that these equations are perfectly general in that they don't appear to require the use of the Butterworth model to be correct. As a matter of fact, we will find that these same equations apply in the cases of the Chebyshev and elliptic filters, with the difference being that the coefficients are derived in different ways.

2.4.2.2 The Chebyshev filter

The design of a Chebyshev filter requires the specification of the maximum ripple in the passband, and the minimum attenuation at some frequency greater than the cutoff frequency. The order of the filter will be derived from these considerations. Once the order is found, the poles of the transfer function are found from the major and minor axes of the ellipse:

The radius of the minor axis is $a\omega_c$ where

$$a = \frac{1}{2}\left(\gamma^{\frac{1}{n}} - \gamma^{-\frac{1}{n}}\right) \qquad \text{and} \qquad \gamma = \frac{1}{\varepsilon} + \sqrt{1 + 1/\varepsilon^2} \qquad (2\text{-}42)$$

and the radius of the major axis is $b_{\omega c}$ where

$$b = \frac{1}{2}\left(\gamma^{\frac{1}{n}} + \gamma^{-\frac{1}{n}}\right) \qquad (2\text{-}43)$$

Thus a and b are functions of known quantities, the maximum permissible ripple given by ε, and the order of the filter, n.

Since there is no ripple in first order filters, it makes little sense to talk about first order Chebyshev filters. Nevertheless, the transfer function for such a filter would be

$$H(s) = \frac{1}{1 + \varepsilon s/s_i} \qquad (2\text{-}44)$$

where the only pole is clearly on the real axis of the s-plane. There are no zeros, which is the same condition as exists in the Butterworth filter family.

The cutoff frequency for the transfer function in Equation (2-44) is

$$f_c = \frac{1}{2\pi}\frac{s_i}{\varepsilon} \qquad (2\text{-}45)$$

Now if we were to calculate the transfer function for the sampled analog filter based on this Chebyshev model by the standard z transform approach, we would make the following calculations:

A. Find the residues, as was done for the Butterworth filter,

B. Take the inverse LaPlace transform

C. Sample the impulse response

D. Take the Z transform of the sampled impulse response.

The result for first order is

$$H(z) = \frac{1}{1 - e^{-\frac{s_i}{\varepsilon}T} z^{-1}} \qquad (2\text{-}46)$$

Given in this form, it looks like what was encountered in Section 2.1, except in this case, the exponential in the denominator has, in its argument, the maximum ripple factor ε.

The sampled analog cutoff frequency turns out to be

$$f_c = \frac{1}{2\pi T} \cos^{-1} \left(\frac{1 - 4b_1 + b_1^2}{2b_1} \right) \qquad (2\text{-}47)$$

which is the same equation as we had in Section 2.1.

The same Chebyshev filter could be transformed by the bilinear z method with a resulting different set of sampled analog coefficients. Making the usual substitution

$$s = \frac{2}{T} \frac{z - 1}{z + 1}$$

we get a transfer function which is very similar to the transfer function of Equation 2.4 except that the maximum ripple factor, ε, is once again included in the denominator.

The coefficients are as follows:

$$a_o = 1 \qquad b_o = 1$$

$$a_1 = 1 \qquad b_1 = \frac{s_i T - 2\varepsilon}{s_i T + 2\varepsilon}$$

Because of the requirement for stability, the coefficient, b_1, is limited to absolute values between zero and one. For these values, the cutoff frequency can be found. The equation for the sampled analog cutoff frequency is given in Equation (2-5) in Section 2.1. It is repeated here for clarity:

$$f_c = \frac{1}{2\pi T} \cos^{-1} \left(\frac{-2b_1}{1 + b_1^2}\right) \qquad (2\text{-}5)$$

As before, this equation can be solved for b_1 in terms of f_c

$$b_1 = \frac{-1 \pm \sin 2\pi f_c T}{\cos 2\pi f_c T} \qquad (2\text{-}6)$$

The second order Chebyshev filter does exhibit ripple. The passband contains a ripple no greater than ε, and the stopband is monotonic. The second order Chebyshev polynomial is

$$P_2(s) = 2s^2 - 1 \qquad (2\text{-}48)$$

The form of the squared magnitude of the filter is

$$|H_2(s)|^2 = \frac{1}{1 + \varepsilon^2 P_n^2(s)} \qquad (2\text{-}49)$$

where $P_o = 1$, $P_1 = s$, $P_2 = 2s^2 - 1$ and, in general, $P_{n+1} = 2sP_n - P_{n-1}$. Often P is given in terms of s/s_i.

In order to design a second order Chebyshev sampled analog filter, it is necessary, as always, to find the inverse LaPlace transform of the $H_2(s)$ transfer function (for the impulse invariant method). This means finding the residues at the two poles of the transfer function. The poles are found in the same way as in the first order with the additional complexity being that the poles are complex, not real. Each pole lies on the ellipse defined earlier in this section. The position of the poles on the ellipse is found by dividing the circle of revolution in the s-plane into 2n angular sectors, each with angular measure π/n. If n is odd, then one boundary of the sectors lies on the positive real axis. If n is even, then the positive real axis lies directly in the center of one of the sectors. This is illustrated in Figure 2.44.

The precise position of the poles are then found by the following procedure:

A. The abscissa of the coordinates of the pole is specified by the point on the circle defined by the minor axis of the ellipse, and through which the boundary of the sector passes.

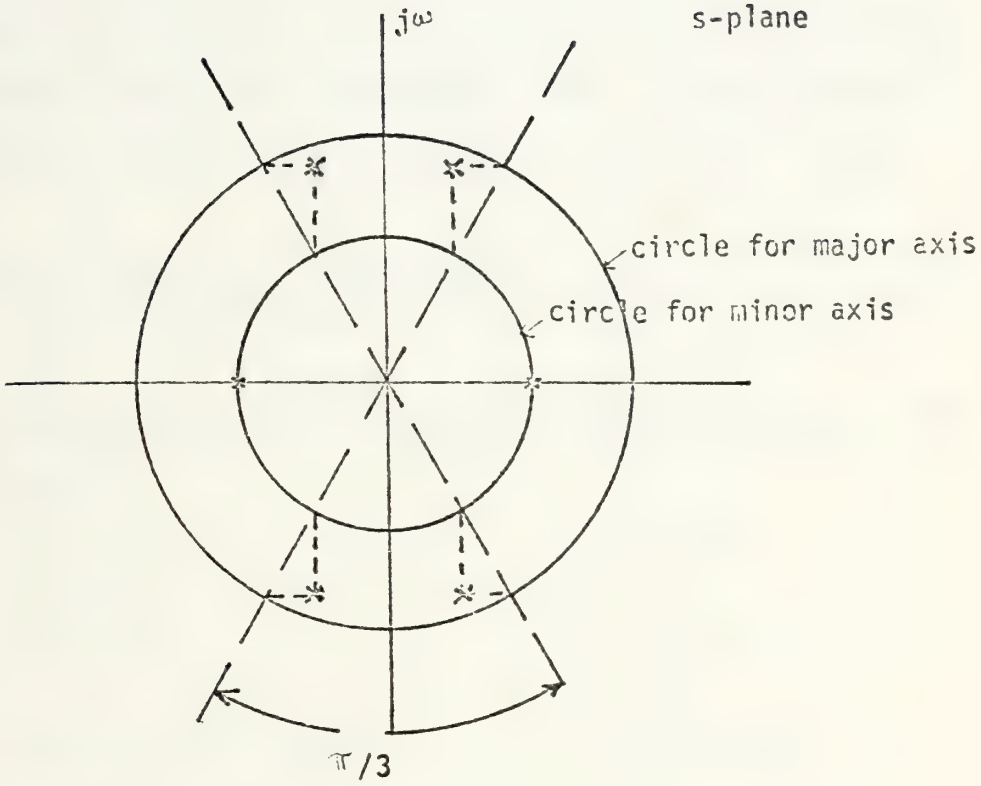


Figure 2.44. Location of poles for a third order Chebyshev filter.

B. The ordinate of the coordinates of the pole is specified by the point on the circle defined by the major axis of the ellipse, and through which the boundary of the sector passes.

After the poles are found, the transfer function must be arranged in the residue form, inverse LaPlace transform taken, sampled, and the Z transform taken. This procedure (now standard) results in the following transfer function:

$$Z(h(n)) = \frac{2/\beta\ e^{-\alpha T} \sin \beta T\ z^{-1}}{1 - 2e^{-\alpha T}\cos \beta T\, z^{-1} + e^{-2\alpha T} z^{-2}} \qquad (2\text{-}49)$$

and the sampled analog coefficients can be easily seen from the above form:

$$a_o = 0 \qquad b_o = 1$$

$$a_1 = 2/\beta\ e^{-\alpha T} \sin \beta T \qquad b_1 = -2e^{-\alpha T} \cos \beta T$$

$$a_2 = 0 \qquad b_2 = e^{-2\alpha T}$$

where α and β are the real and imaginary parts of the singularities in the s-plane.

The cutoff frequency for this analog filter can be found by determining for which frequency the squared magnitude function is down by 3 db. This entails solving the following equation:

$$1 + \varepsilon^2 (2(\frac{\omega_c}{\omega_i})^2 - 1)^2 = 2(1 + \varepsilon^2) \qquad (2\text{-}50)$$

which, of course, yields four roots. For any real ε, the equation will yield one positive root, one negative root, and two purely imaginary roots. Naturally, it is the positive real root which we want for our cutoff frequency. For $\varepsilon = 0.01$, for instance, the squared roots of the equation are found by

$$(\frac{\omega_c}{\omega_i})^2 = \frac{1 + \ 2 + 1/\varepsilon^2}{2} \qquad (2\text{-}51)$$

which yields roots at +50.5 and -49.5. The square roots of these numbers are approximately ± 7.1 and ±j7.0. Thus the cutoff frequency is about 7.1 ω_i.

For the sampled analog case, the cutoff frequency can be found in the same manner. The resulting expression is identical to that found in the section on Butterworth filters:

$$f_c = \frac{1}{2\pi T} \cos^{-1}\{\frac{\cos\ \beta T(1+e^{-2\alpha T}) \pm \sqrt{\cos^2 \beta T(1+2e^{-4\alpha T}+e^{-6\alpha T})}}{2e^{-2\alpha T}} \ \ldots$$

$$\ldots + \frac{\cos\beta T[-8e^{-\alpha T} + 4e^{-2\alpha T} - 8e^{-3\alpha T}]+1+10e^{-2\alpha T}+e^{-4\alpha T}}{2e^{-2\alpha T}}\} \qquad (2\text{-}52)$$

The quantities α and β are, of course, different for the Butterworth and Chebyshev cases.

In order to use the bilinear z transformation, we need not determine residues or find the inverse LaPlace transform of the transfer function. We make the usual substitution and find that the second order Chebyshev low pass filter in the z-plane has the following transfer function:

$$H(z) = \frac{G(1+2z^{-1}+z^{-2})}{1 + \frac{(-4+2\alpha^2+2\beta^2)}{4+4\alpha+\alpha^2+\beta^2} z^{-1} + \frac{(4-4\alpha+\alpha^2+\beta^2)}{4+4\alpha+\alpha^2+\beta^2} z^{-2}} \qquad (2\text{-}53)$$

and the coefficients are

$$a_o = 1 \qquad b_o = 1$$

$$a_1 = 2 \qquad b_1 = \frac{-4+2\alpha^2+2\beta^2}{4+4\alpha+\alpha^2+\beta^2} \qquad (2\text{-}54)$$

$$a_2 = 1 \qquad b_2 = \frac{4-4\alpha+\alpha^2+\beta^2}{4+4\alpha+\alpha^2+\beta^2}$$

as found before in the Butterworth section.

The cutoff frequency is found in the usual manner. The same expression is obtained as in the Butterworth section. The expression is a quadratic in cos $(2\pi\ f_c/f_s)$ with the following coefficients:

term	coefficient
$\cos^2(2\pi\ f_c/f_s)$	$1 + b_1^2 + b_2^2 + 2b_1 + 2b_1b_2 - 6b_2$
$\cos^1(2\pi\ f_c/f_s)$	$2 + 2b_1^2 + 2b_2^2 + 4b_2$
$\cos^o(2\pi\ f_c/f_s)$	$-1 - b_1^2 - b_2^2 + 2b_1b_2 + 2b_1 + 6b_2$

(2-55)

Once again this is a complicated expression, but given the poles, it is a straightforward procedure to find, first the coefficients, and then f_c.

2.4.2.3 The Elliptic Filter

The elliptic filter (Cauer parameter filter) is a natural outgrowth of the Chebyshev filter. The Chebyshev filter is characterized by having ripples in the passband and a monotonic characteristic in the stopband. The elliptic filter "spreads" the error in the stopband of a Chebyshev filter by producing ripples in the stopband as well as in the passband. The elliptic filter utilizing this procedure produces the best (shortest) transition region for any filter of a given order and given ripple amplitude. In other words, the elliptic filter produces the steepest skirts from passband to stopband.

The specifications of an elliptic filter must begin with the maximum ripple in the passband which is defined by ε. The maximum ripple in the stopband is given by $1/A$. In addition, the cutoff frequency, f_c, and the cutin frequency, f_r, must be specified.

The squared magnitude form of the filter is given by

$$|H(s)|^2 = \frac{1}{1 + \varepsilon^2 F^2(s)} \qquad (2\text{-}55)$$

where F is an elliptic function.

The elliptic function, sometimes written as sn(u,k) must be solved to find each pole and zero location. The expression sn(u,k) is shown below where u is the independent variable and k is a parameter. If y = sn(u,k) then

$$u = \int_0^y \frac{dt}{(1-t^2)^{1/2}(1-k^2t^2)^{1/2}} = sn^{-1}y \qquad (2\text{-}56)$$

The form of F can be understood by the following two forms:

If the filter is to be of even order, then

$$F(\omega) = \frac{(\omega_1^2-\omega^2)(\omega_2^2-\omega^2) \ \ldots \ (\omega_{2n-1}^2-\omega^2)}{(1-\omega^2\omega_1^2)(1-\omega^2\omega_2^2) \ \ldots \ (1-\omega^2\omega_{2n-1}^2)} \qquad (2\text{-}57)$$

If the filter is to be of odd order, then

$$F(\omega) = \frac{(\omega_2^2-\omega^2)(\omega_4^2-\omega^2) \ \ldots \ (\omega_{2n}^2-\omega^2)}{(1-\omega_2^2\omega^2)(1-\omega_4^2\omega^2) \ \ldots \ (1-\omega_{2n}^2\omega^2)} \qquad (2\text{-}58)$$

The trick is to choose the zeros of F (and thus, the poles since if there is a zero at ω_1, then there is a pole at $1/\omega_1$) such that the maximum ripple ε, the minimum attenuation A, and the f_c and f_r are fulfilled. To do this requires solving the incomplete elliptic integral shown in Equation (2-56).

In order to calculate the coefficients of a sampled analog filter, by the impulse invariant method, it is first necessary to find the poles of the analog transfer function. The poles must be found by solving

$$1 + \varepsilon^2 sn^2(u,k) = 0 \qquad (2\text{-}59)$$

No rigorous attempt has been made to solve this equation. Values for various arguments of sn(u,k) have been calculated and are available in tables.

After the poles are found, the derivation of the sampled analog filter types continues along the same lines as for the Butterworth and Chebyshev filters with the single exception that the analog filter form contains zeros. This fact does not appear to be significant, however, for purposes of design in the z-plane.

2.4.3 DIRECT SELECTION OF COEFFICIENTS FOR THE SAMPLED ANALOG SECOND ORDER RECURSIVE FILTER

The derivation of a sampled-analog filter from well known analog filter types is attractive because some of the qualities of the analog filters are attractive. Unfortunately, upon transformation, many of their attractive traits are lost. Certainly, by use of the standard z transform, one does not even have the assurance that the ranges of the analog filter and the sampled analog filter are the same. On the other hand, the frequency "warping" effect spoken of in Section 2.1 which occurs when the bilinear z transformation is used often is a hindrance to proper design.

It appeared, in our research, that what was really needed by the design engineer was some method to determine, first, the admissible values of, say, b_1 and b_2 in a sampled analog

filter; second, the general shape of the transfer function as a function of b_1 and b_2; and third, some indication as to how much critical parameters change when the coefficients are shifted slightly due to temperature variations, aging, etc. This last consideration is especially important since the recursive filter has been held in some disfavor by designers because of possible stability problems, alleged problems of unstable parameters even if the filter itself is stable, and finally, because of its infinite duration unit pulse response (time domain). If it could be shown that the filter of a certain design was stable and had stable parameters over a specified tolerance of coefficients, then the idea of a recursive filter in systems should be encouraged.

To that end, an effort was made to discover the admissible values of b_1 and b_2 in a second order sampled analog recursive filter, whose transfer function is as follows:

$$H(z) = \frac{a_o + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}} \qquad (2\text{-}59)$$

There are three criteria defining admissibility of a point whose coordinates are (b_2, b_1). One criterion is the stability constraint, a physical constraint. Thus, if z_1 is a pole of the system, then

$$|z_1| < 1 \qquad (2\text{-}60)$$

The other two criteria defining admissibility arise from basic algebra. For the quadratic equation

$$z^2 + b_1 z + b_2 = 0$$

where z_1 and z_2 are roots to the equation, then

$$z_1 z_2 = b_2 \tag{2-61}$$

and

$$z_1 + z_2 = - b_1 \tag{2-62}$$

For purposes of clarity, it is appropriate to treat the admissibility problem in three distinct cases. For the first case, we set b_2 negative.

A. CASE ONE: b_2 negative:

If b_2 is negative, then the polarity of the poles (their signs) are opposite. Using the third criterion (Equation (2-62)), we observe that the absolute value of b_1 must be less than the larger of the absolute values of z_1 and z_2. A more useful observation for this case is that there can never be complex roots to the polynomial. The discriminant of a quadratic equation, $ax^2 + bx + c = 0$ is

$$b^2 - 4ac$$

and, in our case, the discriminant is

$$b_1^2 - 4b_2$$

Thus, if b_2 is negative, then the discriminant is positive for all values of both b_1 and b_2, thus insuring that each root is real.

A graph can clearly show the first two criteria: Figure 2.45 shows the z_1 and z_2 axes with a hyperbola

$$z_1 z_2 = b_2 \quad (b_2 \text{ is negative}) \tag{2-63}$$

Also superimposed onto the graph is the unit square. Since neither pole can have an absolute value greater than 1, the poles must lie within the square. Combining those two criteria, we have the darkened portion of the hyperbola drawn in Figure 2.45 which is the locus of the only available pole locations. Figure 2.46 shows the admissible values of b_1 based on the data in Figure 2.45. (If $z_1 = 1$, then $z_2 = -b_2$ and $b_1 = -(1-b_2)$ etc.)

Figure 2.47 shows the locus of admissible pole locations on the z-plane for this case. Note that one pole varies from b_2 to 1 while the other varies from -1 to $-b_2$. Finally, Figure 2.48 shows the important result: the permissible values of b_1 with respect to b_2 for b_2 negative. The diagonal lines have the following equations:

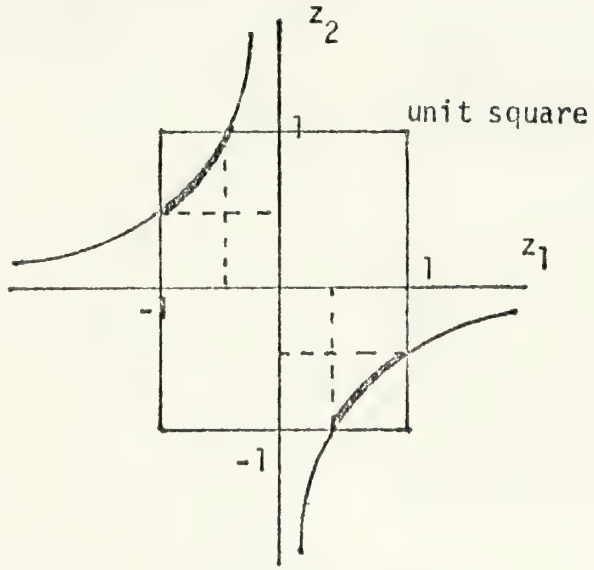


Figure 2.45. Admissible pole locations for $b_2 < 0$

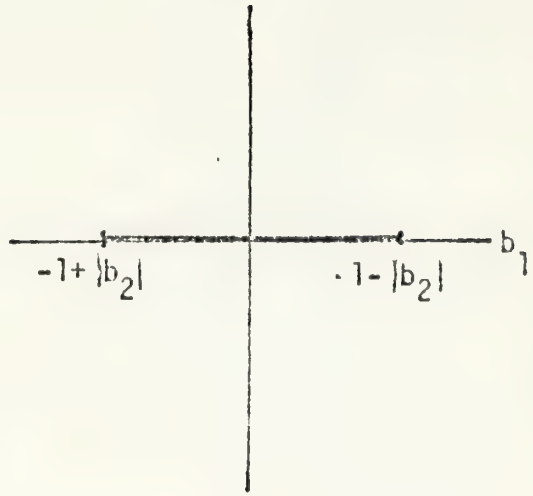


Figure 2.46. Admissible values of b_1 for $b_2 < 0$

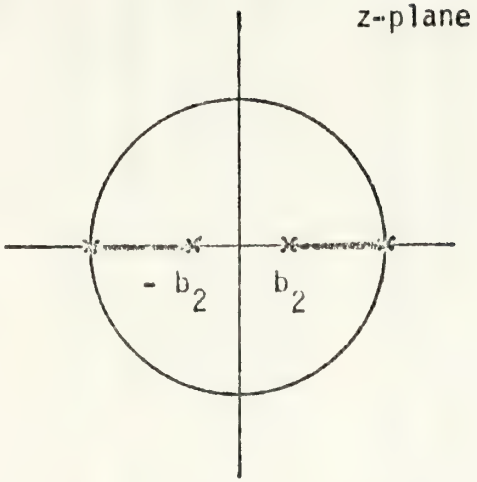


Figure 2.47. Loci of admissible pole locations for $b_2 < 0$

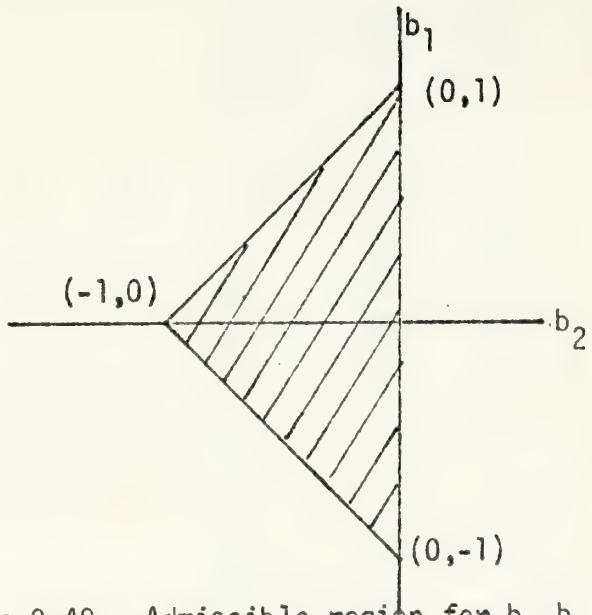


Figure 2.48. Admissible region for b_1, b_2 $b_2 < 0$

$$b_1 = 1 + b_2 \qquad \text{(upper line)} \tag{2-64}$$

$$b_1 = -1 - b_2 \qquad \text{(lower line)} \tag{2-65}$$

B. CASE TWO: b_2 is positive and the poles are real.

We initially note that the relationship between the poles and b_1 is

$$z_1 + z_2 = - b_1 \tag{2-62}$$

This time, since b_2 is positive, each pole must have the same sign. Thus the absolute value of b_1 is always greater than either z_1 or z_2. We also note that, in this case, the discriminant is not necessarily positive. Nevertheless we shall enforce this condition artificially and take up complex roots in CASE THREE.

The hyperbola is once again shown in Figure 2.49. The darkened segments of the hyperbola indicate the admissible values for the poles given some positive b_2. Figure 2.50 shows the admissible values of b_1 based on the admissible locations of the poles.

Going further, Figure 2.51 shows the admissible pole locations on the z-plane and Figure 2.52 extends Figure 2.48 to show the admissible values of b_1 given a positive b_2 and real roots.

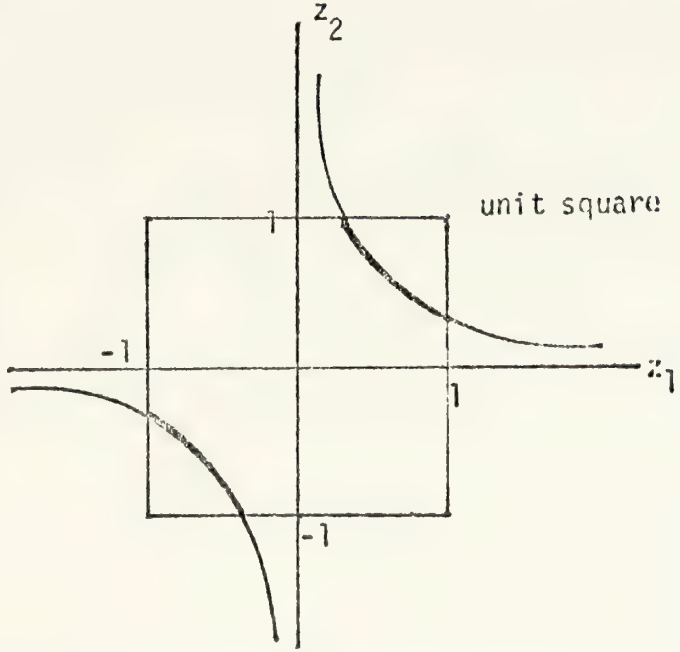


Figure 2.49. Admissible pole locations
$b_2 > 0$

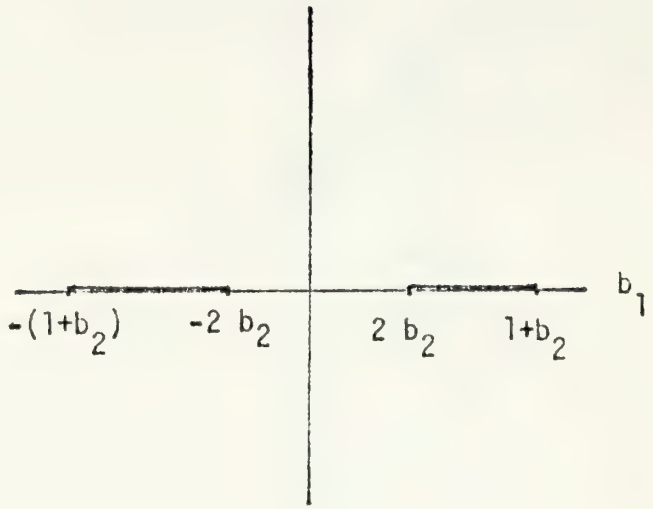


Figure 2.50. Admissible values of b_1
$b_2 > 0$

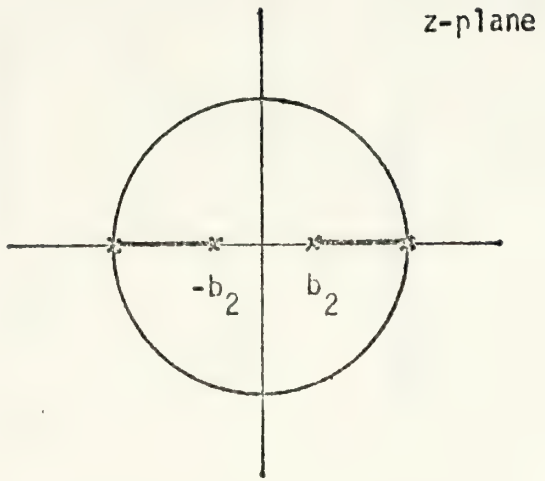


Figure 2.51. Loci of admissible pole locations, $b_2 > 0$

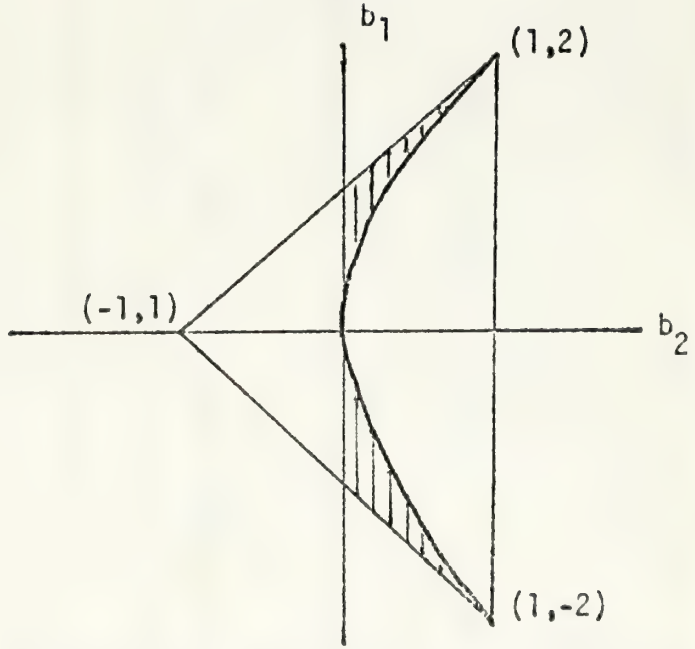


Figure 2.52. Admissible region for b_1, b_2 $b_2 > 0$, and real roots

The region is bounded by the extension of the two lines given in Equations (2-64) and (2-65) plus the horizontal parabola

$$b_1 = 2\sqrt{b_2} \tag{2-66}$$

C. CASE THREE: b_2 is positive and the poles are complex.

For complex poles, the job is made easier by the observation that b_2 is always the square of the absolute value of each of the two complex conjugate poles: If $z_1 = x + jy$ and $z_2 = x - jy$, then the absolute value of either of the poles is

$$\sqrt{x^2 + y^2}$$

Applying the algebraic criteria, we have

$$b_1 = -2x$$

$$b_2 = x^2 + y^2$$

which, as stated, is the square of the absolute value of either pole. Since we know that the absolute value of each pole cannot exceed one, we can limit our search to values of b_2 less than 1.

Figure 2.53 shows a root locus diagram (same as Figure 2.51 except extended to include complex roots). As the pole locations are moved around in the b_2 circle, b_1 changes considerably. As a final graph, Figure 2.54 shows the total admissible region for b_2 and b_1 including all three cases discussed here. Notice that the region is convex (the complex case having filled in the formerly concave region) and is bounded by three line segments:

$$b_1 = 1 + b_2 \tag{2-64}$$

$$b_1 = -1 - b_2 \tag{2-65}$$

$$b_2 = 1 \tag{2-67}$$

The full meaning of Figure 2.54 can be realized if an engineer is visualized in the laboratory attempting to design some second order recursive filter. He cannot choose his b_1 and b_2 at random. He must choose them so that the point (b_2,b_1) lies within the region shown in Figure 2.54. We have labeled this region the b-PLANE TRIANGLE OF STABILITY since at all points on and outside the triangle, the filter becomes unstable.

It remains to be determined how the triangle itself can be partitioned. Certianly, the triangle gives us the region of real roots and of complex roots, but their significance cannot be grasped completely without further partitioning.

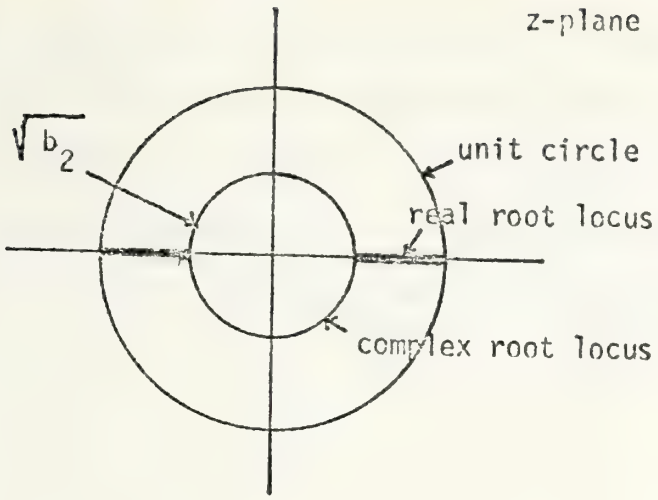


Figure 2.53. Loci of admissible pole locations

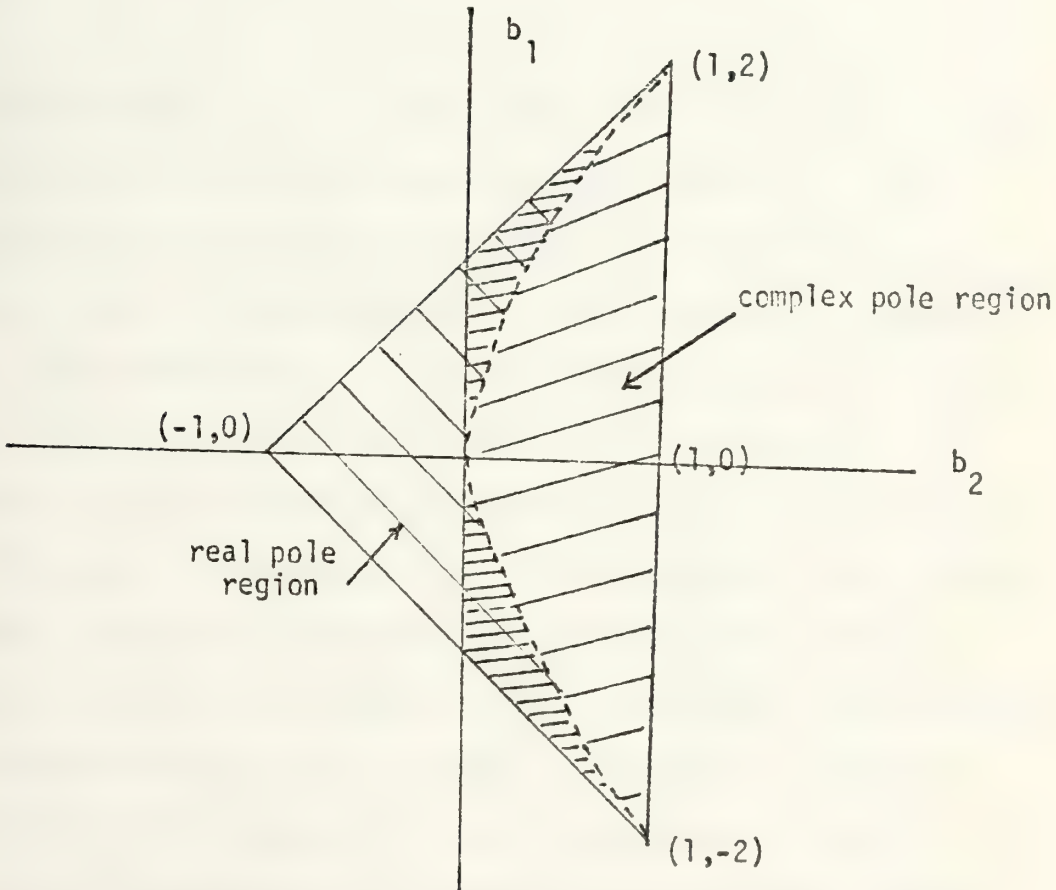


Figure 2.54. Admissible region for b_1, b_2

It was with this in mind that we began to investigate a filter's characteristic based on a region in the b-plane rather than on the pole-zero plot on the z-plane. For this test, we held the coefficients a_o, a_1, and a_2 at the following values: (those naturally arising from the bilinear z transform)

$$a_o = 1$$

$$a_1 = \pm 2$$

$$a_2 = 1$$

These values of a's insure that zeros will exist in the transfer function of the sampled analog filter. This is important for the radar processing application we had in mind and which is discussed in greater detail in Chapter three. The condition $a_1 = +2$ indicates a low pass filter (second order zero at $z = -1$) and the condition $a_1 = -2$ indicates a high pass filter (second order zero at $z = 1$).

A computer program was written to calculate the cutoff frequency for many different combinations of b_2 and b_1, insuring that the stability criterion was met. This program is labeled STACY15 and is listed in Appendix D. Figure 2.55 contains the results of these programs. Figure 2.55 shows cutoff frequency as a fraction of sampling frequency for low pass filters, while Figure 2.56 shows the cutoff frequency as a fraction of sampling frequency for high pass

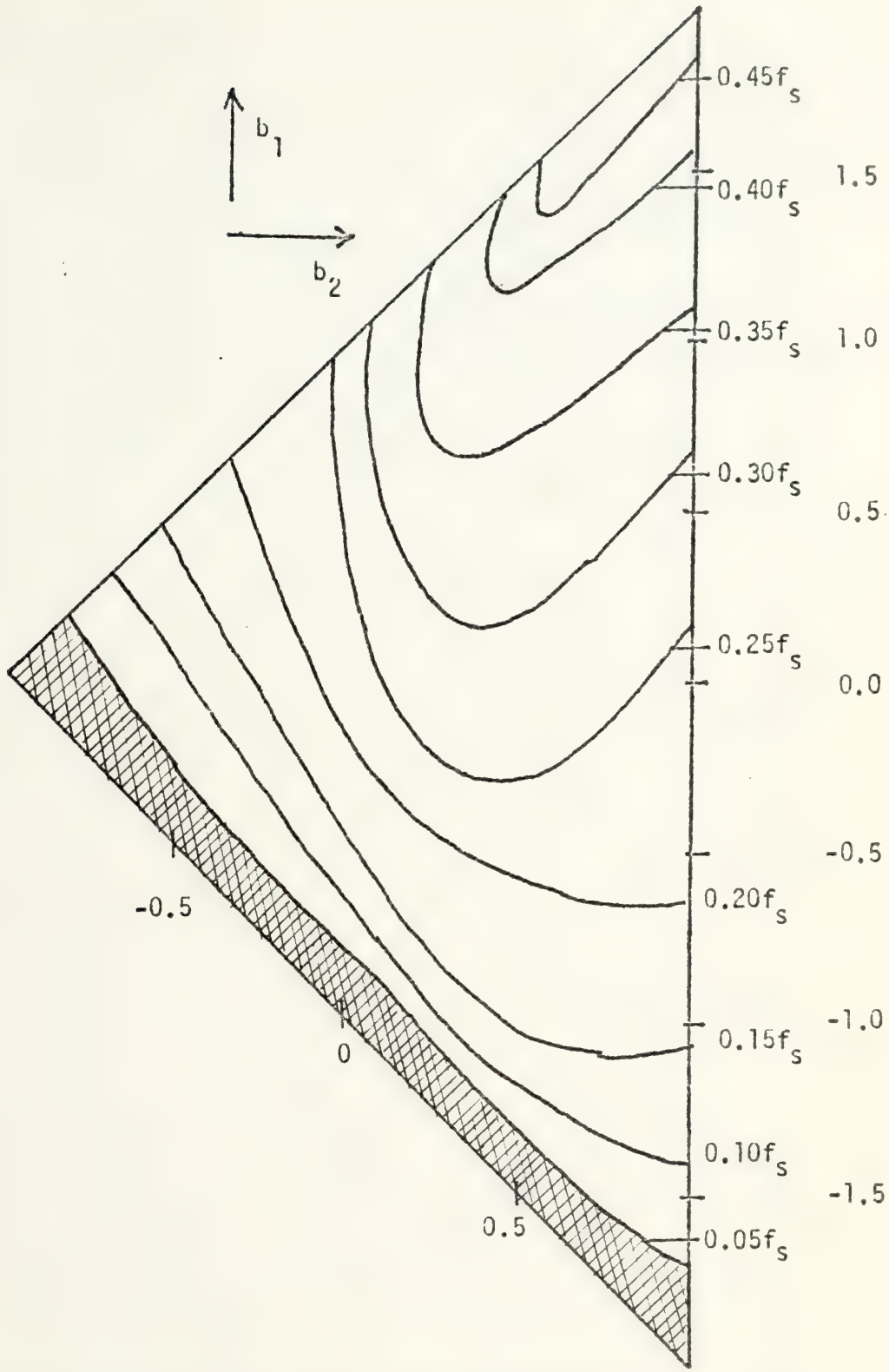


Figure 2.55. Isocontours for cutoff frequency of a second order low pass filter. $a_0=1$, $a_1=2$, $a_2=1$. The b-plane triangle of stability.

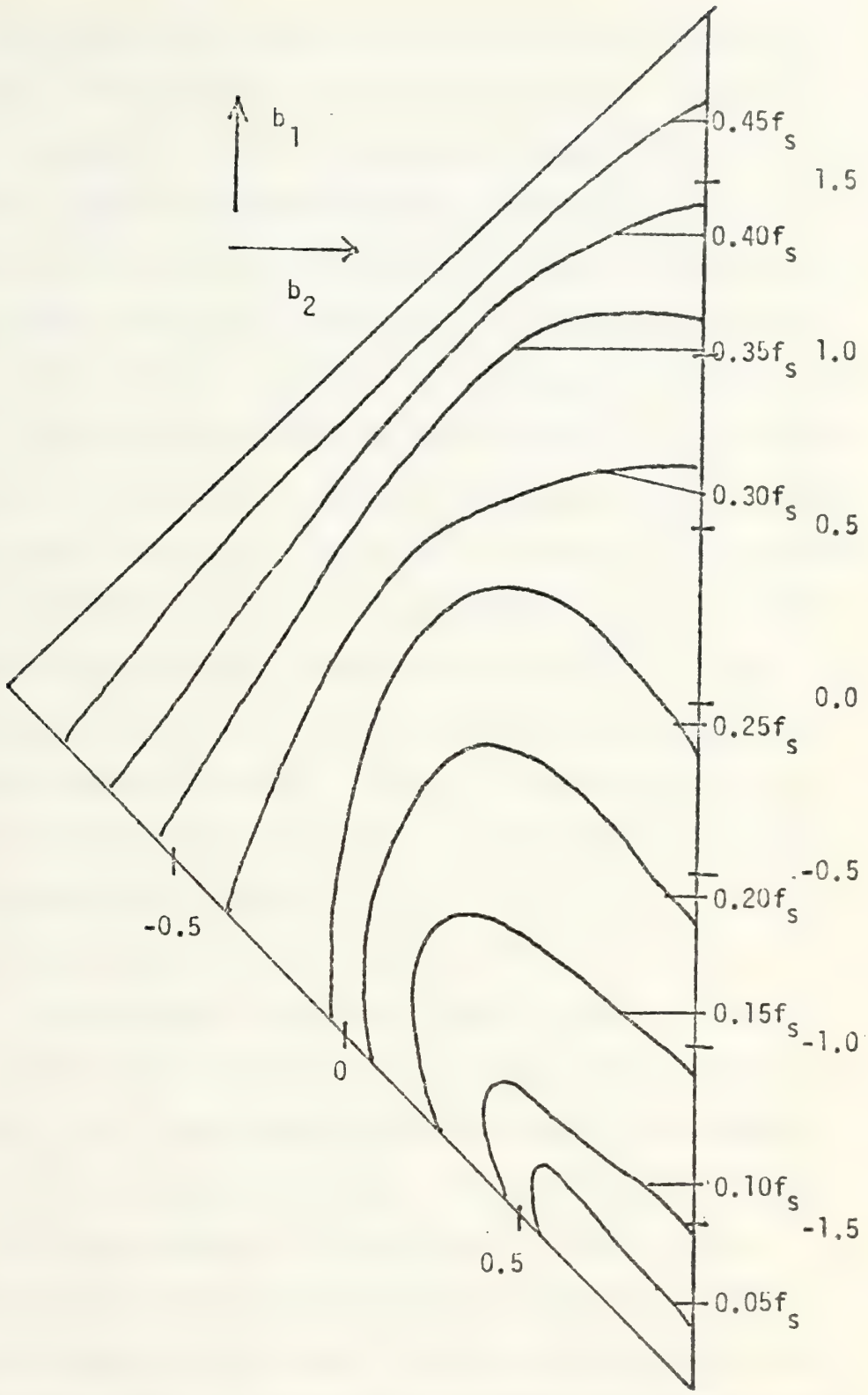


Figure 2.56. Isocontours for cutoff frequency of a second order high pass filter. $a_0=1, a_1=-2, a_2=1$. The b-plane triangle of stability.

filters. The triangle of stability has been divided into ten sections. Each section consists of b-pairs such that the cutoff frequencies lie in a certain range. For instance, for extremely narrow low pass filters, we would want our coordinates to lie in the crosshatched region of Figure 2.55, for this is where the cutoff frequency is less than $0.05f_s$.

Each solid line is an "iso-cutoff frequency line" ... a line whose locus of points defines filters with equal cutoff frequencies. Of particular interest is the solid line whose cutoff frequency is $0.25f_s$. This is the dividing line between integrators (below the line) and cancellers (above the line).

As can be seen from the triangle of stability, large latitude for the selection of b_1 and b_2 exists when seeking a narrow band low pass filter. On the other hand, only a small region exists for a canceller of equivalent Q (cutoff frequency greater than $0.45f_s$).

Several other observations may be made:

A. When complex roots are used, there is, in general, a maximum of the transfer function somewhere between dc and $f_s/2$. When real roots are used, the maximum exists at dc (for low pass filters) or $f_s/2$ (for high pass filters). The consequence of this result is that if we require our filter to be monotonic in the passband, then, generally, we require real roots. Such a conclusion is not completely valid, since there is a small region where even with complex roots, the transfer function is monotonic in

both passband and stopband. Now real roots work pretty well for integrator type second order filters, but less well for canceller types. The narrow region between the complex-real roots dividing line and the boundary of the triangle for canceller type filters probably precludes the design of a useful second order canceller (low pass) using real roots. The integrator type has no such limitations as can be seen on Figure 2.56.

B. Rough estimates of how the cutoff frequency varies with b_1 and b_2 can also be seen from Figures 2.55 and 2.56. Consider Figure 2.55 at about $b_2 = 0.15$. The cutoff frequency is $0.3f_s$, and hardly changes at all for values of b_1 from 0.5 to 1.2. Thus we would say that the directional derivative in the b_1 direction for the cutoff frequency in this region is zero.

All the contours have these regions of nearly constant cutoff frequency.

C. Some contours show marked insensitivity to changes in b_2 as well. Consider the $0.25f_s$ contour. For $b_1 = -0.4$, it is shown that the cutoff frequency hardly changes for values of b_2 from 0.5 to 0.75.

D. The direction of increasing f_c for Figure 2.55 is toward the upper right hand corner of the b-plane triangle.

Figure 2.57 (ripple) shows the dependence of ripple on values of (b_2,b_1). As can be seen, the entire region of the triangle where real roots are found is in a region of zero

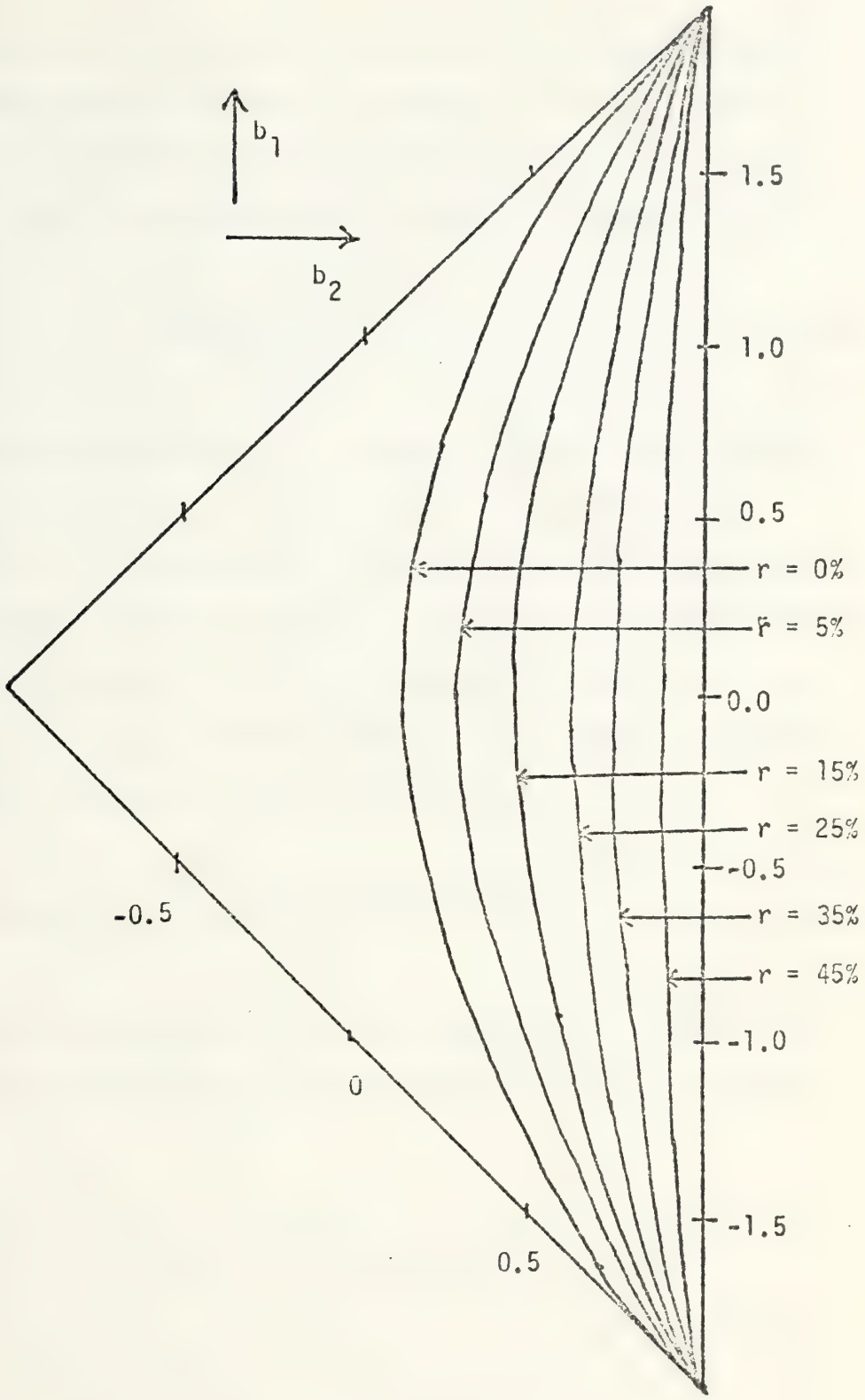


Figure 2.57. Isocontours for ripple (r) of a second order low pass filter. $a_0=1$, $a_1=2$, $a_2=1$. The b-plane triangle of stability

ripple. Moreover, the zero ripple region extends into the complex root region as shown in Figure 2.57 (the region between the complex-real root contour and the zero ripple contour). For this calculation, ripple is defined by

$$r = \frac{|H_{max}(f)| - |H(0)|}{2|H_{max}(f)|} \tag{2-68}$$

Such a definition permits, at times, certain undesirable mathematical results. For instance, for exceptionally large ripple, it is possible that the amplitude response at the cutoff frequency, $|H(f_c)|$, is actually larger than the response at zero, $|H(0)|$. To see how this arises, we need to have a precise definition of the cutoff frequency used in this research. That definition is

$$|H(f_c)| = \frac{.707\ |H_{max}(f)| + |H(0)|}{2} \tag{2-69}$$

which uses, as a basis, an average value of the amplitude response in the passband. Now whenever $|H(f_c)| = |H(0)|$ we have

$$2|H(f_c)| = .707\ (|H_{max}(f)| + |H(0)|)$$

or

$$1.293\ |H(f_c)| = .707\ |H_{max}(f)|$$

and

$$|H(0)| = .547\ |H_{max}(f)|$$

so that the ripple is

$$r = \frac{|H_{max}(f)| - 0.547\ |H_{max}(f)|}{2\ |H_{max}(f)|} = .227 \qquad (2\text{-}70)$$

Low pass filters with a ripple of 22.7% or more are, in general, less useful than filters with lower ripple. Thus, it would appear that this awkward situation only occurs in filters which are not particularly useful.

It should be noted, at this point, that the ripple figures for the high pass filter have been omitted. The reason for this is that the ripple contours are in exactly the same place as in the low pass filter. Of course, the definition of ripple, for high pass filters, is modified to

$$r = \frac{|H_{max}(f)| - |H(f_s/2)|}{2\ |H_{max}(f)|} \qquad (2\text{-}71)$$

But the resulting contours are identical to the ones shown in Figure 2.57.

The next figure (2.58) shows the iso-contours relating to the important quantity, "skirt slope". In analog filters, a well known rule of thumb holds that the skirt of the filter (its transition region between passband and stopband) falls 6 db per octave for each real pole. Thus, this

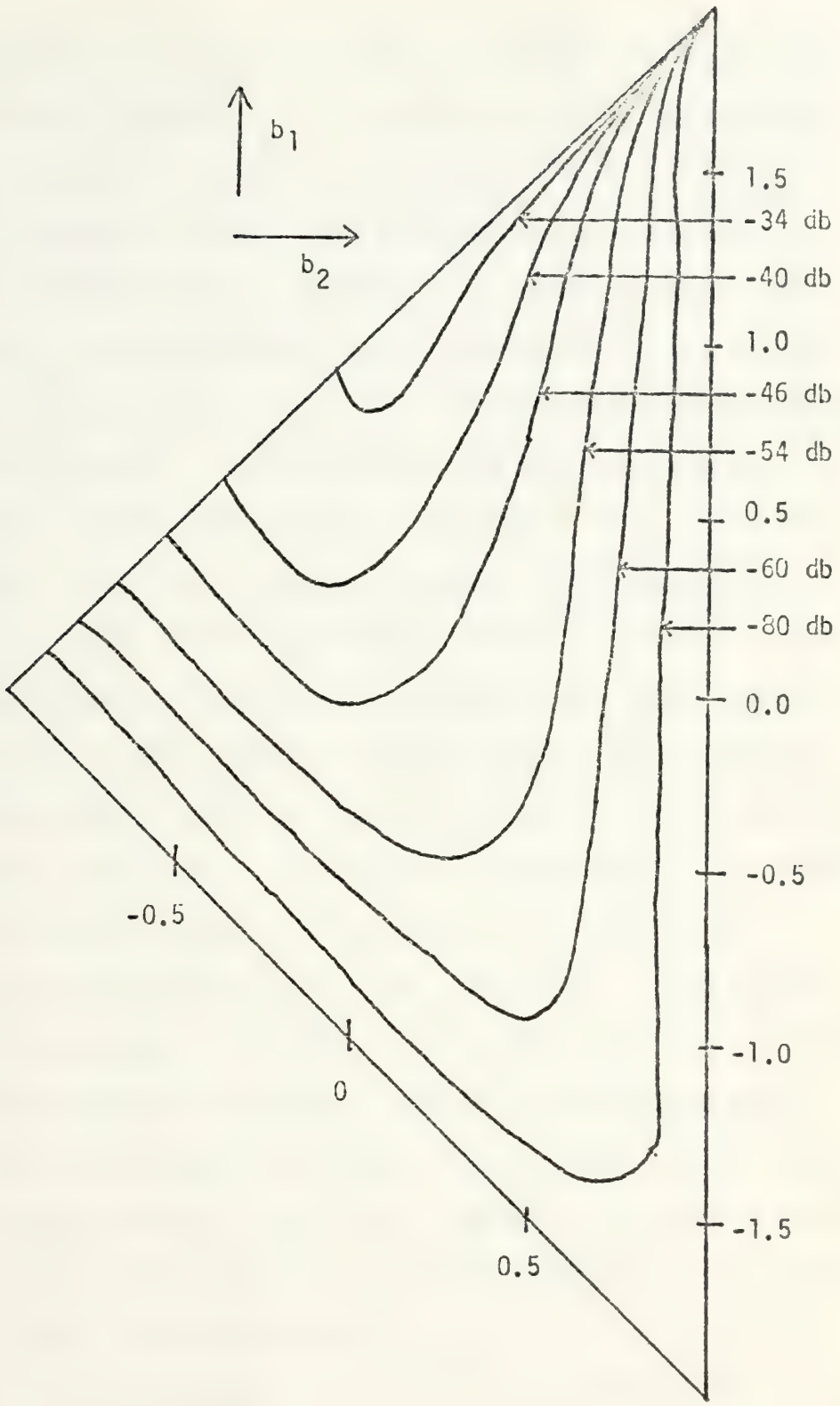


Figure 2.58. Isocontours for the slope of the transfer function at cutoff for a second order low pass filter. $a_0=1$, $a_1=2$, $a_2=1$. The b-plane triangle of stability. Slopes are given in terms of db per decade.

specification is often the determining factor for the order of Butterworth, Chebyshev, or elliptic filter. As can be seen in the figure, no such easy rule of thumb is possible in sampled analog filters. The slopes of the transfer function at cutoff are, of course, all negative (for low pass filters). The greatest slope (absolute value) occurs for filters whose (b_2,b_1) pair is located in the lower right part of the triangle. This is also the region in which (Figure 2.55) cutoff frequency is extremely low. That is, these filters have a very narrow passband and extremely large slopes. The smallest slope contour is for -5. This contour cuts through many cutoff frequencies, from approximately 0.23 f_s to the highest cutoff frequencies. Filters of this nature have lower Q and are generally not so useful as the higher Q filters. If ripple is considered, we observe that for a maximum of 5% ripple, we have a wide range of possible slopes to choose from, or, we have a large range of cutoff frequencies from which to select. However, selecting a cutoff frequency severely limits the possible skirt slopes. As can be seen by a brief study of the first three figures in this section (2.55, 57, and 58), a selection of, say, $f_c = 0.15f_s$ and less than 5% ripple limits us to slopes between -51 and -38 db/decade.

As an aid to locating the frequency for which the transfer function is a maximum, Figure 2.59 is included. The curves are once again equi-contours such that they are the

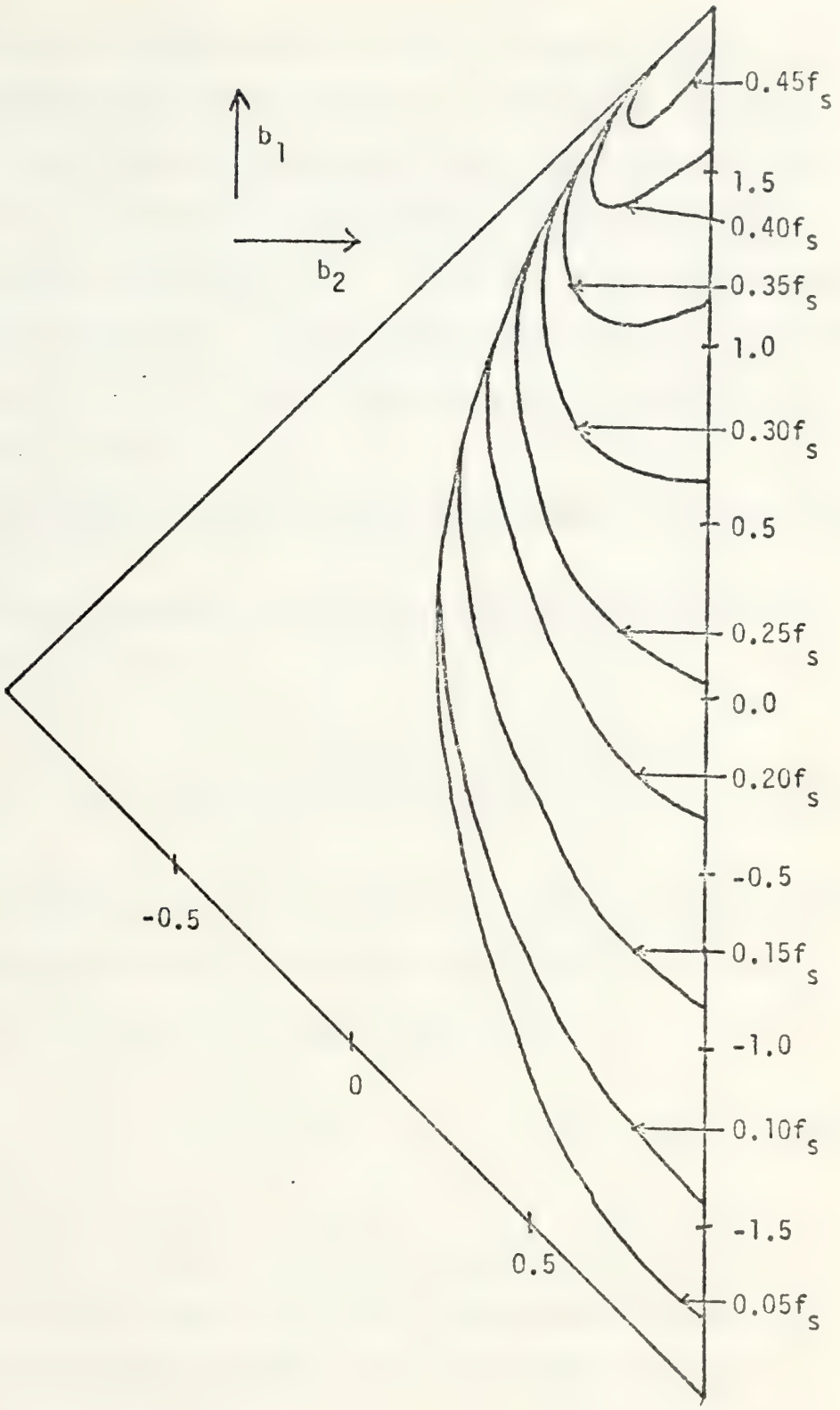


Figure 2.59. Isocontours for the position of the maximum of the transfer function for a low pass second order filter. $a_o=1$, $a_1=2$, $a_2=1$. The b-plane triangle of stability.

locus of points for which the maximum of the transfer function occurs at a given frequency. Naturally, if the ripple is zero, then the maximum of the transfer function occurs at f = 0. Thus, the zero ripple contour is also the contour for which $f_{max} = 0$. Since all the contours are restricted to the region of non-zero ripple, the information on the figure is not so significant unless a designer is forced to use ripple.

AN EXAMPLE OF DESIGN USING THE CHARTS (Figures 2.55, 57,58,59)

What is the frequency characteristic of the following transfer function in z?

$$H(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - 0.8z^{-1} + 0.5z^{-2}} \qquad (2\text{-}72)$$

Finding the frequency characteristic can be done graphically using the distances from zeros and poles, or can be done analytically by making the substitution

$$z = e^{j2\pi f/f_s} \qquad (2\text{-}73)$$

and solving the equation for various values of f.

It can also be done, to a first approximation, by referring to the four previous figures. The following procedure is recommended:

A. Locate the position of $H_{max}(f)$ from Figure 2.59 ... In this case, $f_{max} = 0.13f_s$. Place a point on a graph

with the coordinates (.13f_s, 1) which normalizes the maximum value of the transfer function to one.

B. Consult the ripple chart (Figure 2.57). Determine the ripple. In this case, r = 15%. Recalling the equation for ripple (Equation (2-68)), we then note that H(0) must be 0.7. Place this value on the graph.

C. Since there is a minimum (zero) at $f_s/2$, place a third point at the coordinates (0.5f_s, 0).

D. Now locate the cutoff frequency from Figure 2.55. In this case, $f_c = 0.20f_s$. Recalling the definiiton of cutoff frequency (Equation (2-69), we have $H(f_c) = 0.6$. Place the point (0.2f_s, 0.6) on the graph.

E. Refer to the slope chart (Figure 2.58). The slope at cutoff frequency is about -57 db/decade. This is fairly large. Draw a tangent line through the transfer function at cutoff frequency with approximately a -57 db/decade slope.

F. Finally, connect the points with a smooth curve. There will never be unexpected ripples in the frequency characteristic. The frequency characteristic will be monotonically increasing from dc to the frequency for which H is a maximum, and will then be monotonically decreasing from H_{max} to $H(f_s/2)$.

Figures 2.60 through 2.63 show the procedures for determining the approximate shape of the frequency transfer function.

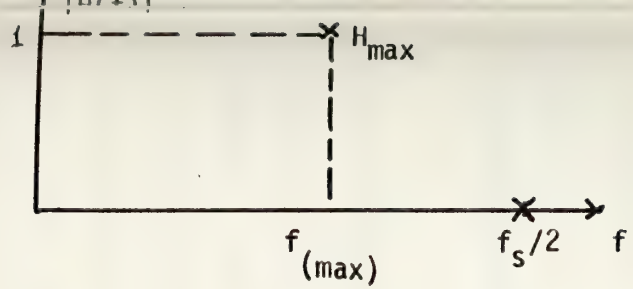


Figure 2.60. Location of H_{max} from figure 2.59. Also, $H(f_s/2) = 0$.

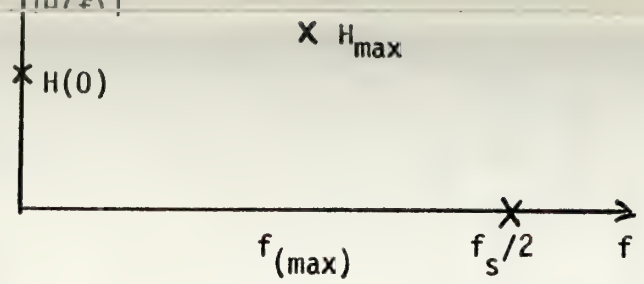


Figure 2.61. Calculation of H(0) by determining ripple from figure 2.57.

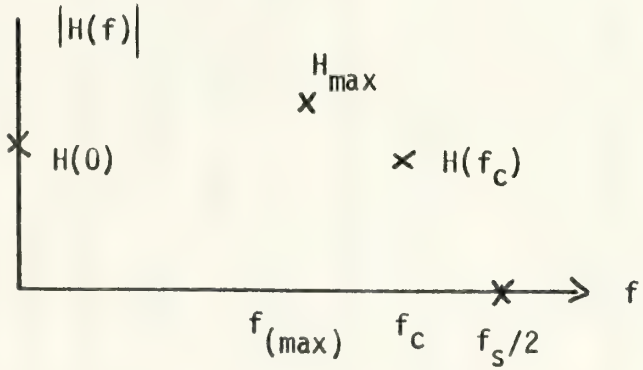


Figure 2.62. Location of cutoff frequency from figure 2.55

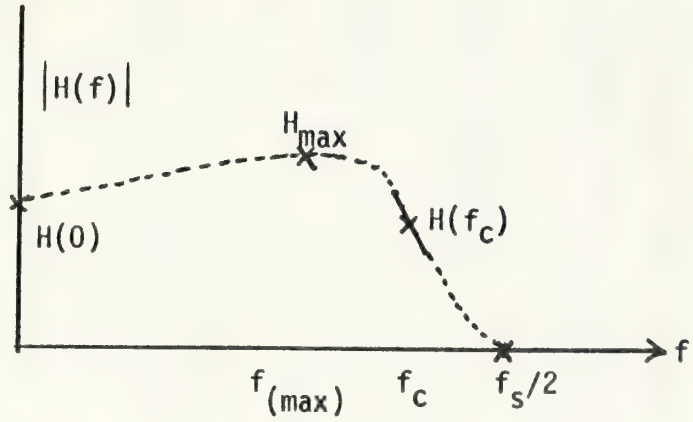


Figure 2.63. Determination of slope at cutoff from figure 2.58 and connecting the points.

A SECOND EXAMPLE:

Given a required specification (skirt slope, cutoff frequency, and ripple), determine the appropriate values for b_1 and b_2.

The specifications to be used are

$$f_c = 0.2f_s$$

ripple < 5%

skirt slope --- maximum, but greater than 60 db/decade

The following procedure is recommended:

A. Locate the locus of points for which $f_c = 0.2f_s$. It is shown on Figure 2.55.

B. Intersect this locus with the locus of points for which the ripple = 5%. All points on the cutoff frequency locus to the left of the intersection meet the first two specifications.

C. Consult Figure 2.58, the skirt slope chart. The -60 db/decade contour is not attainable with such a small ripple. The best that can be done is about -32 db/decade. This occurs at $b_1 = -0.62$ and $b_2 = 0.36$. By actually going through this example, one can see that the skirt slope specification can be met, but only by changing one of the other two specifications. For lower cutoff frequencies, say $f_c = 0.1f_s$, slopes of -40 db/decade can easily be attained at no higher ripple.

The use of recursive sampled analog filters using CCD's has been limited because of the stability problem and the anticipation that the parameters of the CCD may not be sufficiently constant to retain the desired filter specifications. With this in mind, a sensitivity study was undertaken to determine how cutoff frequency and slope changes with marginal changes in CCD gain. From the canonical form of the second order filter, it can be seen that CCD voltage gain bears an important relationship to the coefficients a_1, a_2, b_1, b_2. Usually these coefficients are implemented with potentiometers, that is, with passive devices whose maximum gain is 1.0. If we desire $b_1 = 0.8$, for instance, we must consider the gain of the CCD in cascade with the potentiometer for b_1. The equation for b_1 is

$$b_1 = (K)(P)$$

where K is the gain of the CCD and P is the potentiometer setting (see Section 2.2.1). If the gain of the CCD were 1.5, then, for this case, we would have

$$P = 0.8/1.5 = 0.533$$

It follows that a 5% change in CCD gain produces a 5% change in b_1.

For the second order coefficients, the gains of each CCD must be considered. The equations for these coefficients are

$$a_2 = K_1 K_2 P \tag{2-74}$$

$$b_2 = K_1 K_2 P \tag{2-75}$$

Once again, it follows that an x% change in CCD gain produces an x% change in the effective coefficients.

Figures 2.64 and 2.65 show the marginal change in cutoff frequency with respect to 5% changes in b_1 and b_2 (each separately). For Figure 2.64, the actual equation used for drawing the contours was

$$f_c = \frac{f_c(1.025b_2) - f_c(0.975b_2)}{f_c(b_2)} \tag{2-76}$$

We observe in this chart that for most of the triangle, a change of 5% in b_2 causes a change of less than 2% in cutoff frequency, which should be tolerable for most applications. The particularly high change regions are near the triangle boundary for negative b_1 and negative b_2. Here the change in cutoff frequency is sometimes greater than 50%. However, a quick look back to Figure 2.55 where the cutoff frequency contours are plotted indicates to us that this is a large change in an extremely small number. Specifically, the region of maximum shift in cutoff frequency with respect

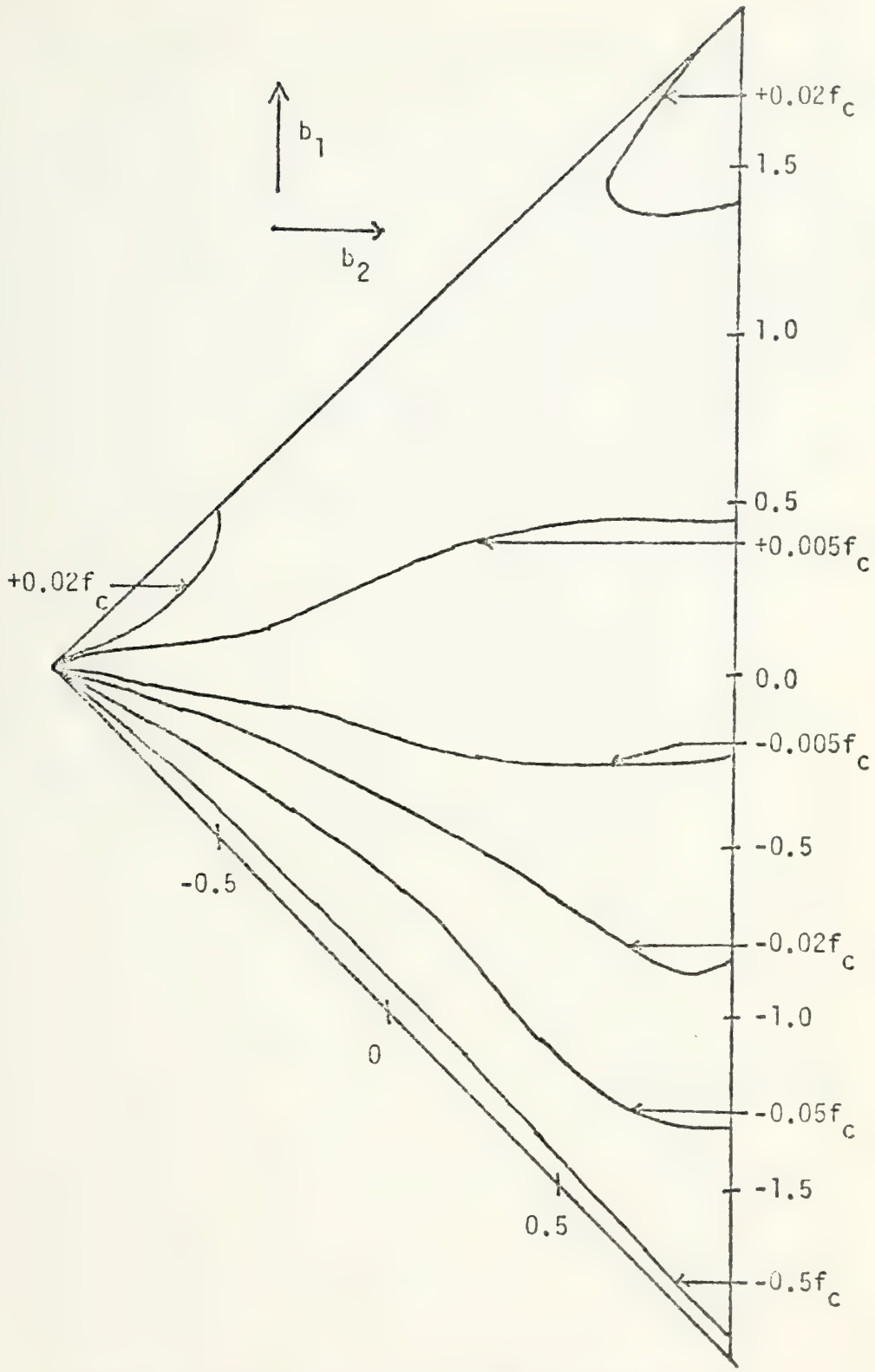


Figure 2.64. Sensitivity of the cutoff frequency as a function of a 5% change in b_1. Isocontours show fractional change in f_c. Low pass, second order filter, $a_0=1$, $a_1=2$, $a_2=1$. δ-plane triangle of stability.

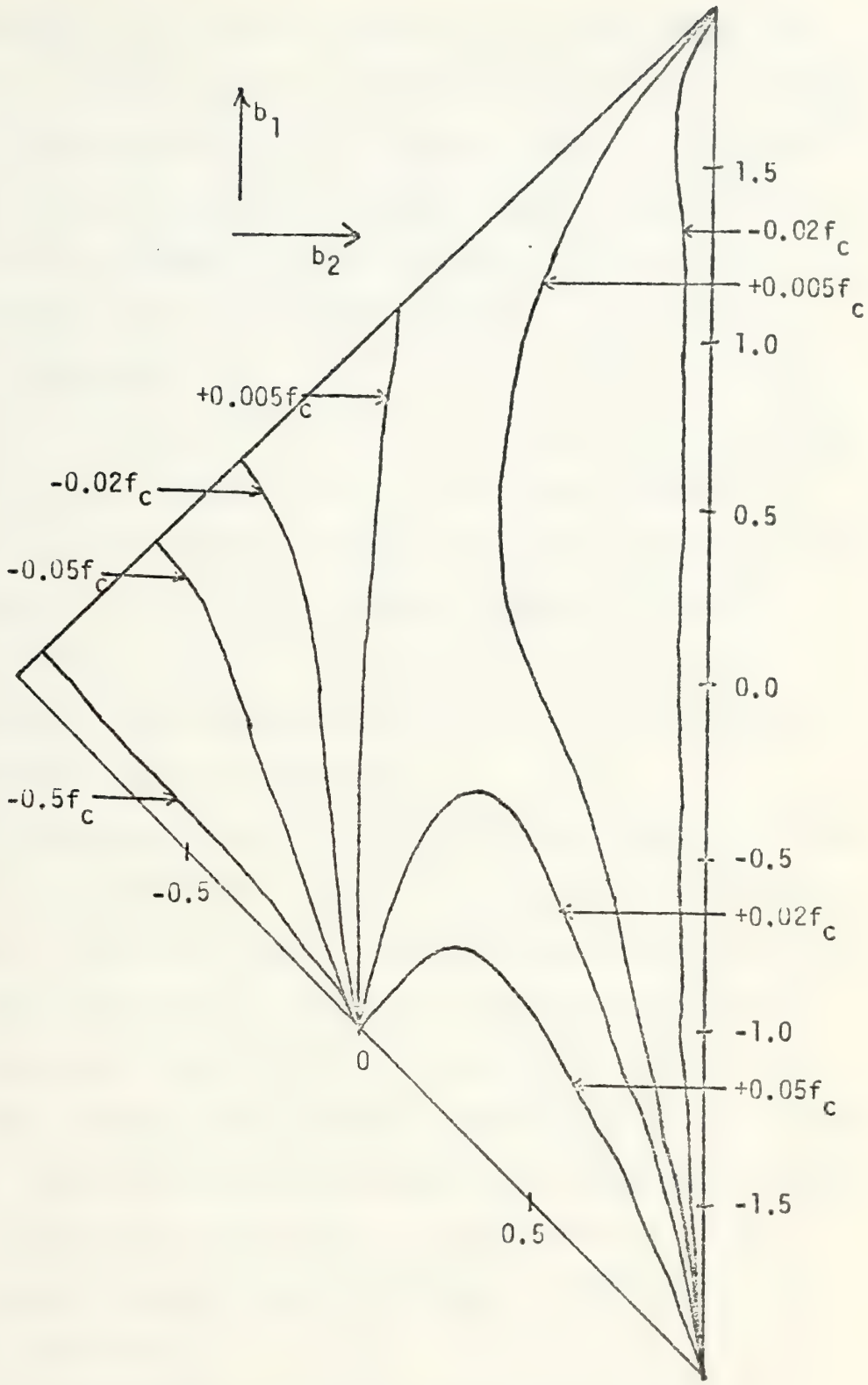


Figure 2.65. Sensitivity of the cutoff frequency as a function of a 5% change in b_2. Isocontours show fractional change in f_c. Low pass, second order filter, $a_0=1$, $a_1=2$, $a_2=1$. b-plane triangle of stability.

to a 5% change in b_2 is the region of smallest cutoff frequency ($f_c < 0.05f_s$). Thus, the absolute shift is probably small, but the relative shift is quite large. This region is probably only of academic interest anyway, since it is so close to the boundary of the triangle. Points close to the boundary are only marginally stable and should be avoided in practice.

Figure 2.65 shows the same shift in cutoff frequency as a function of a 5% change in b_1. As may be seen from an examination of the figure, almost the whole inner region of the triangle above $b_1 = -0.4$ causes a change in cutoff frequency of less than 2%. Once again, the only region where extremely large variations are found is a region of extremely small cutoff frequency.

It is understood that we have calculated the directional derivatives of a function $f_c(b_2,b_1)$ only in the directions of b_2 and b_1. That is, we have not calculated the contours for any other direction of variable change. Intuitively, we could determine the direction of maximum change at each point in the triangle of stability and utilize it for design purposes. However, we believe that such a computation would only be of marginally more use than the present set of curves, and the present curves are easy to visualize for the design engineer.

Example 2 required a given cutoff frequency, ripple, and skirt slope. If we were to include in the specification

that the cutoff frequency could not shift more than, say, 5% with a 5% change in any coefficient, then we must resume our search for the proper coefficients. For the case where $f_c = 0.1f_s$, we find that the additional two charts offer new constraints on where the coefficients can be located. One possible point is $b_1 = -0.9$ and $b_2 = 0.5$. This point meets every constraint except the skirt slope constraint. The slope is only -57 db/decade. If this is not acceptable, and if no other specification can be changed in any way, then either the feedforward coefficients must be changed, or higher order filters must be designed. If the feed forward coefficients are changed, then the zero at $f_s/2$ may be moved or eliminated. For some applications, this is unacceptable.

Finally, two other graphs have been plotted (Figures 2.66, 2.67), each showing the shift in skirt slope with respect to a 5% change in the feedback coefficients. Precisely the same procedure is followed to determine whether a region exists that meets all the specifications of the problem. Briefly, then, the overall procedure is as follows:

A. Using an overlay, trace the admissible regions for each constraint.

B. Observe whether or not there exists any region in the triangle of stability which meets all constraints. If so, select the appropriate coefficients and test.

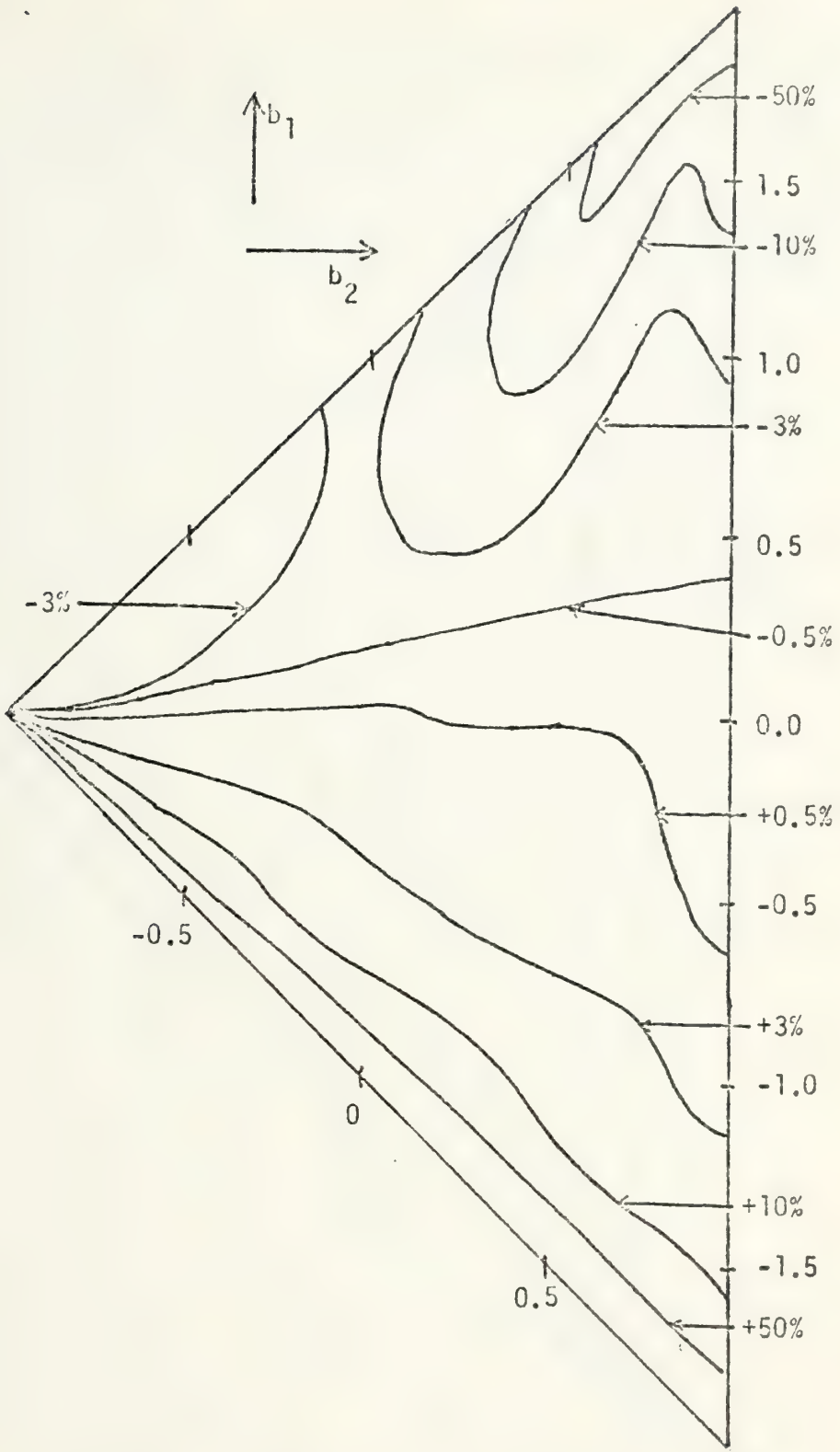


Figure 2.66. Sensitivity of the slope at cutoff as a function of a 5% change in b_1. Isocontours show percentage change in slope. Low pass, second order filter, $a_0=1$, $a_1=2$, $a_2=1$. b-plane triangle of stability.

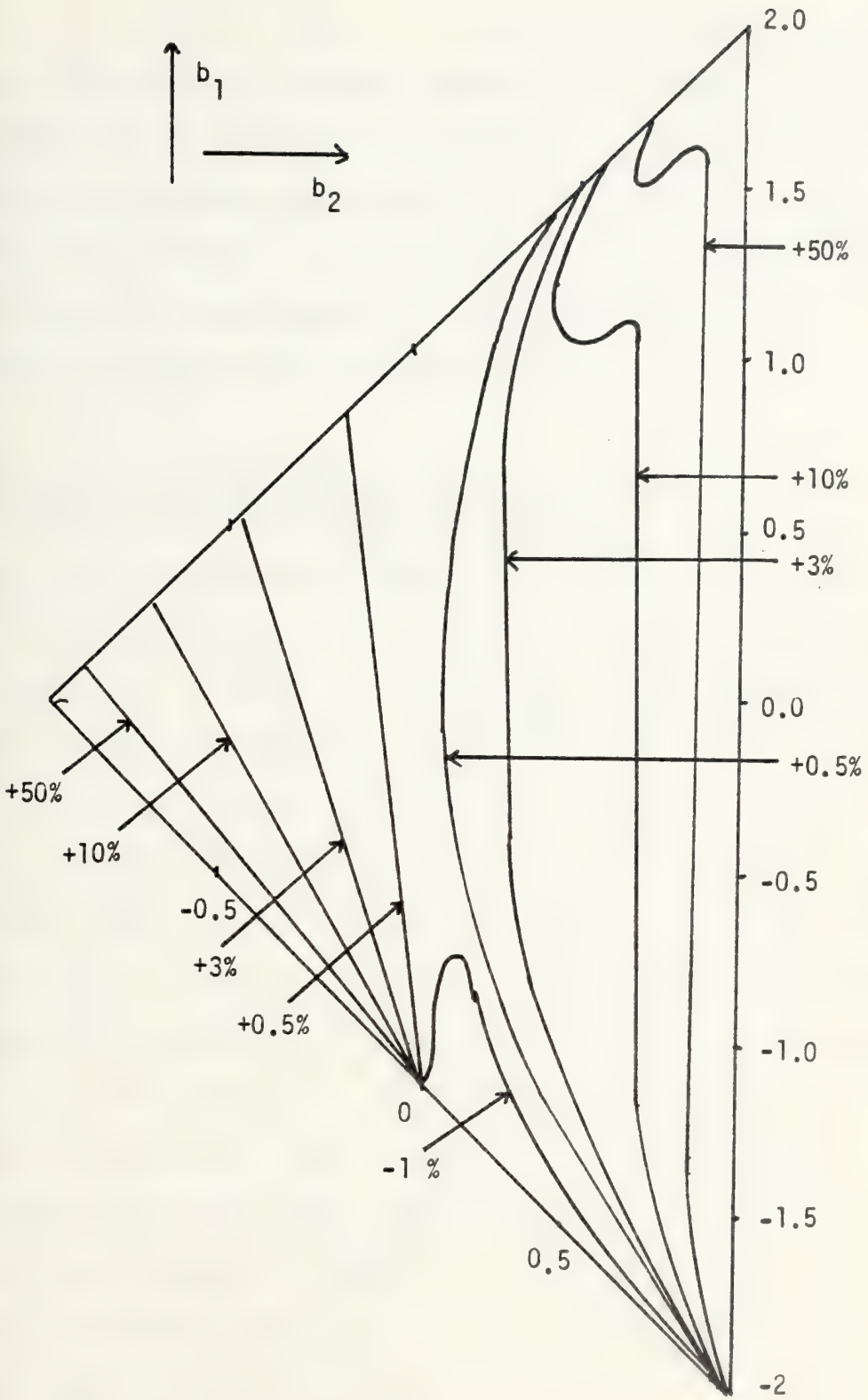


Figure 2.67. Sensitivity of the slope at cutoff as a function of a 5% change in b_2. Isocontours show percentage change in slope. [2]Low pass, second order filter, $a_0 = 1$, $a_1 = 2$, $a_2 = 1$. b-plane triangle of stability.[9]

C. If not, then each constraint must be inspected closely to determine whether relaxation is possible. If not, then one can move the feed forward coefficients (a's), or step up to higher order filters. Moving the a's has a deleterious effect on the zero at $f_s/2$, while additional order filters require more hardware and more study to determine the appropriate coefficients.

2.4.4 DIRECT SAMPLED ANALOG FILTER DESIGN USING COST FUNCTIONS AND COMPUTER ITERATIONS

One common engineering method for determining the values of a multidimensional vector has been the use of cost or loss functions. Briefly, one determines the required ideal function. Then one determines how much the relative cost is for deviation of the actual function from the ideal. This is usually subjectively performed by engineers knowledgeable in the field. A computer program is set up which, given a starting point, proceeds to "step" along the paths of the multi-dimensional vector until a value of the cost function is reduced to an acceptable level, or until a minimum is attained. The result is a set of values for each component in the multi-dimensional vector and a value for the cost function. Commonly, different starting points produce different "convergences" and several starting points must be used in order to insure that a global (or absolute) minimum of the cost function is reached rather than a local minimum. This is especially true for large order problems (where the multi-dimensional vector has many components).

The first to utilize this method for the design of digital filters was Steiglitz in 1970 [16]. His procedure entailed defining a cost function (usually squared error, but not necessarily so) and a desired frequency response from dc to half the sampling frequency (the Nyquist frequency). A starting point is chosen and the computer is used to locate points for the singularities which produce transfer functions which are successively closer to the ideal. When the computer finds a minimum (subject to small rounding errors) it stops and prints out the coefficients, the poles and zero locations, and the squared error.

Using this method, it is not necessary to consult s-plane theory and documentation for an appropriate filter model. Although certain analytical proofs may be lacking (such as the proof that the elliptic filter provides the shortest transition region for any filter of equivalent order), the results are generally so good that no improvement is necessary.

A program was generated as a part of this research to perform the functions detailed above. Modifications to the program have been made to provide for proper weighting regarding the seriousness of deviations from the ideal as a function of frequency. As an example, it is often desired that a perfect zero be present in the stopband. This entails setting the coefficients a_o, a_1, and a_2 to constant values and letting the computer search for the optimum b_1 and b_2. Any number of different combinations can thus be used, each providing different weighting functions and producing a

different filter. The starting point for the direct design of a digital or sampled analog filter then, must be a frequency specification weighted for importance. No other input specification is necessary except for the order of the filter.

Extensive use of the algorithm of steepest descent (Fletcher-Powell algorithm) suggests that several runs be taken for each problem starting at a different point each time. The well known saddle point phenomenon occurs at times and different starting points will sometimes converge to different values. A comparison of the squared error function for each convergence will reveal the true optimum location of the singularities.

One other pitfall is that the Fletcher-Powell algorithm, of course, makes no arbitrary limit on the position of poles and zeros. But our filter must have all its poles inside the unit circle of the z-plane in order to assure stability. Optimum locations of the singularities where the poles are outside the unit circle must be discarded since the filter is unstable. There is a way to eliminate such pole locations however. If we find that z_i is a pole location and $|z_i| > 1$, then we multiply the transfer function by

$$\frac{z - z_i}{z - 1/z_i}$$

which has a magnitude of z_i when z is on the unit circle. This means that the inversion of a pole with respect to the unit circle does not affect the shape of the magnitude characteristic. Since the program is written to find optimum gain as well, such an inversion does not affect the characteristic at all. The program has been written to check automatically to see whether the poles and zeros are outside the unit circle. If they are, they are inverted and the process is begun again at the new locations of poles and zeros.

In addition to being a replacement for the more familiar design by analog models, this process works for any arbitrary frequency characteristic specification. Thus, we are not limited to low pass and high pass models. It is perhaps interesting to note that our frequency characteristic is defined for discrete frequencies of interest. They can, of course, be as close together as desired, but since the computer is digital, the characteristic cannot be totally free of the possibility of unwanted excursions BETWEEN specification points. However, as long as the number of specifications greatly exceeds the order of the filter, this will not be a problem. On the other hand, it can be shown that if the order of the filter is equal to twice the number of frequency specifications, then the specification will, in theory, be met exactly. (This is a result of having 2n degrees of freedom for nth order filters.) However, to meet each frequency characteristic specification precisely

runs the risk of having excessive deviation between the frequency specifications. Actually, this problem is more academic than practical, since in the overwhelming majority of cases, the order is low and the frequency specification is given as a continuous function. It is well to remember that, in the case of a piecewise constant transfer function, more than one frequency, in general, should be included as a specification for each constant region.

Figures 2.68, 2.69, and 2.70 show some of the results of using the Fletcher-Powell-Steiglitz method with the program generated during this research. Figure 2.68 shows a second order low pass filter where the feed forward coefficients were permitted to change. Figure 2.69 shows the same type filter where the feed forward coefficients were held constant at (1,2,1). Figure 2.70 shows the first "tooth" of the filter in Figure 2.69.

Section 2.5 Experimental Results and Theoretical Calculations of Representative Second Order Comb Filters

During the design of the second order comb filter using two CCD's we dealt with several complexities not found in the first order case:

A. Two CCD's had to be found which had similar characteristics in order to "share" bias levels. One of our requirements was that these bias levels be common for corresponding gates on each CCD. We tested and characterized several CCD's in the search for such similarity. We finally

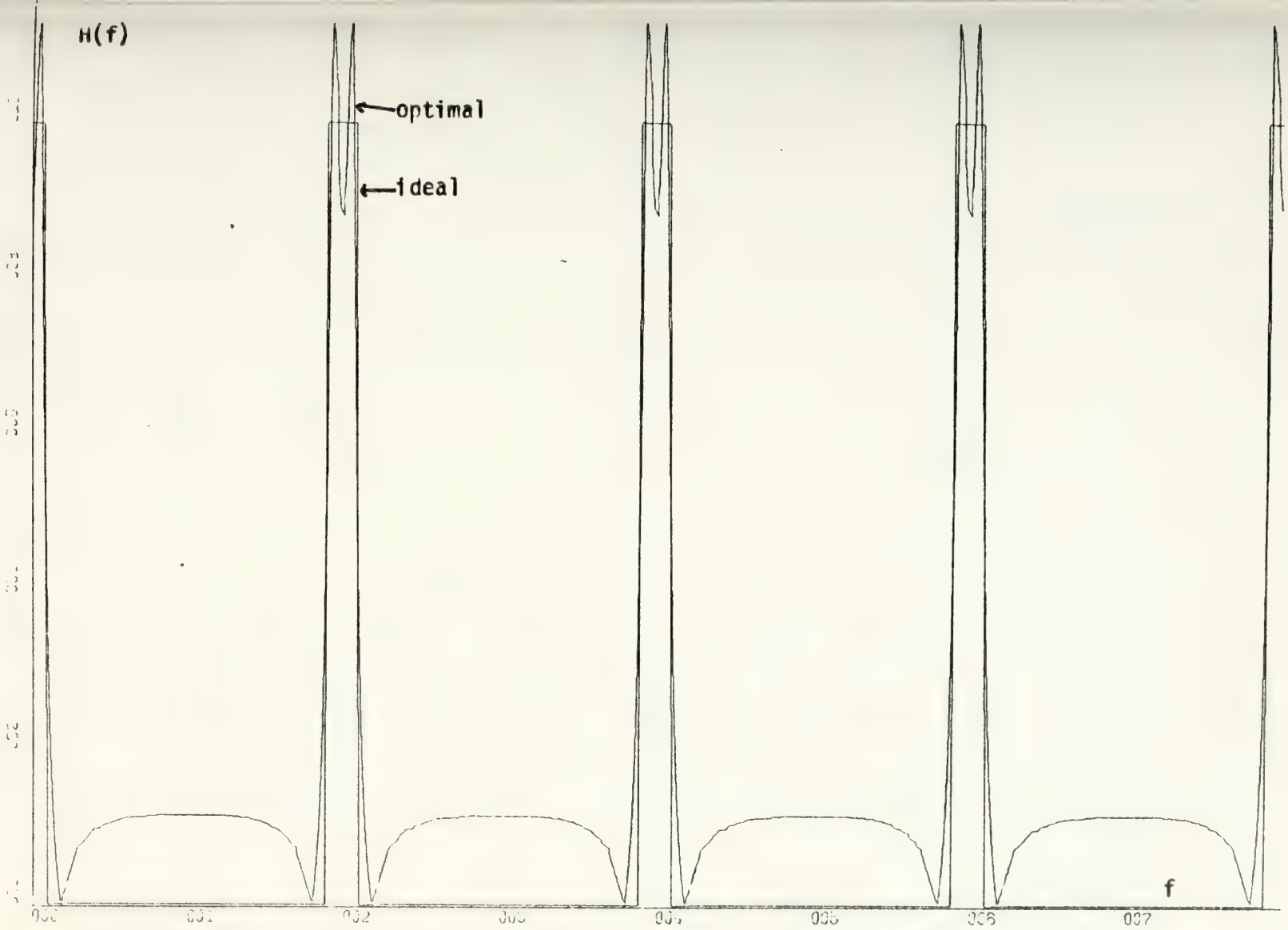


Figure 2.68. Transfer function of a second order, low pass, sampled analog filter where feedforward coefficients were permitted to change.

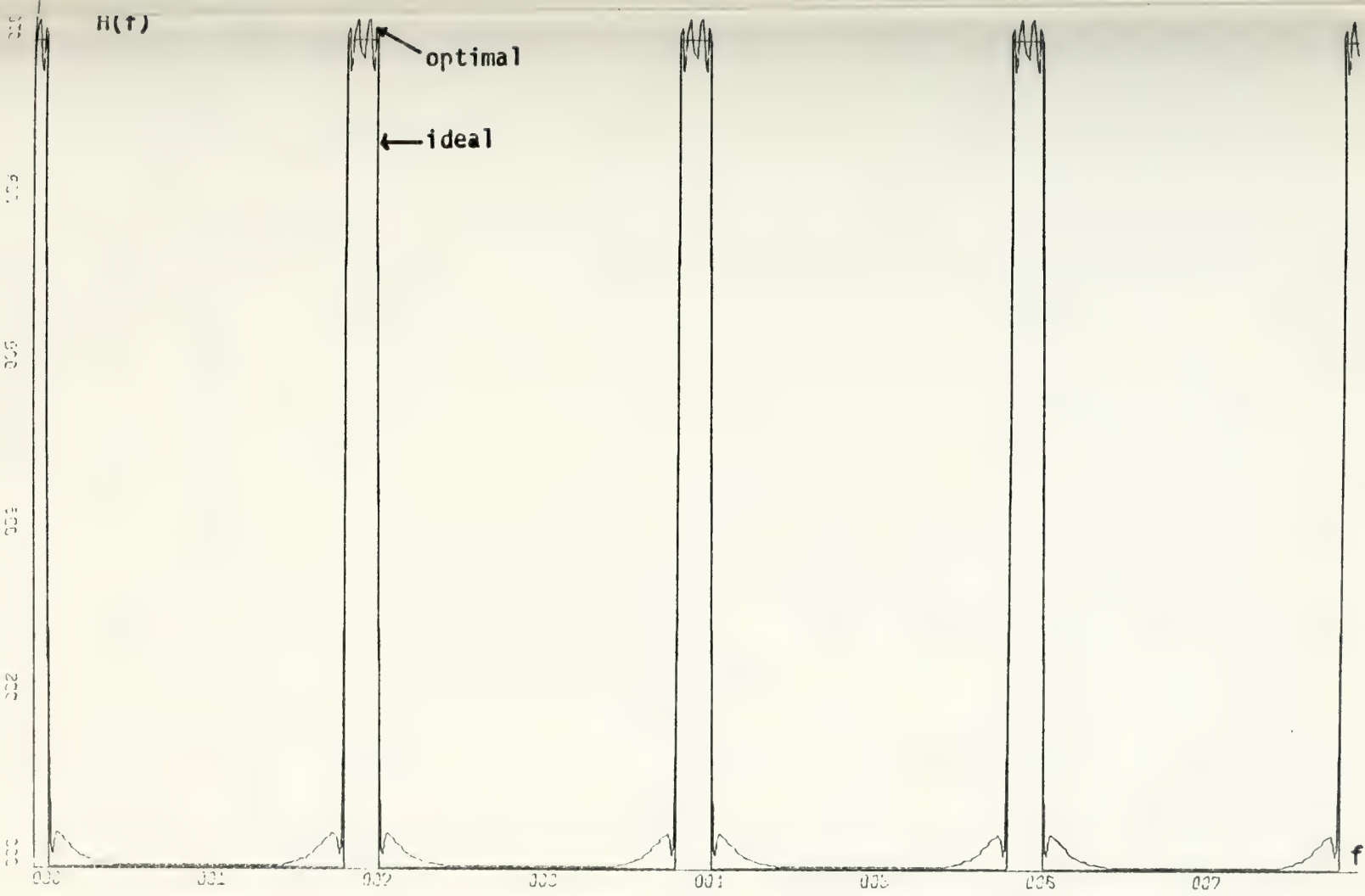


Figure 2.69. Transfer function of a fourth order, low pass, sampled analog filter where feedforward coefficients were held constant at (1,2,1). Notice the zeros at multiples of half the sampling frequency.

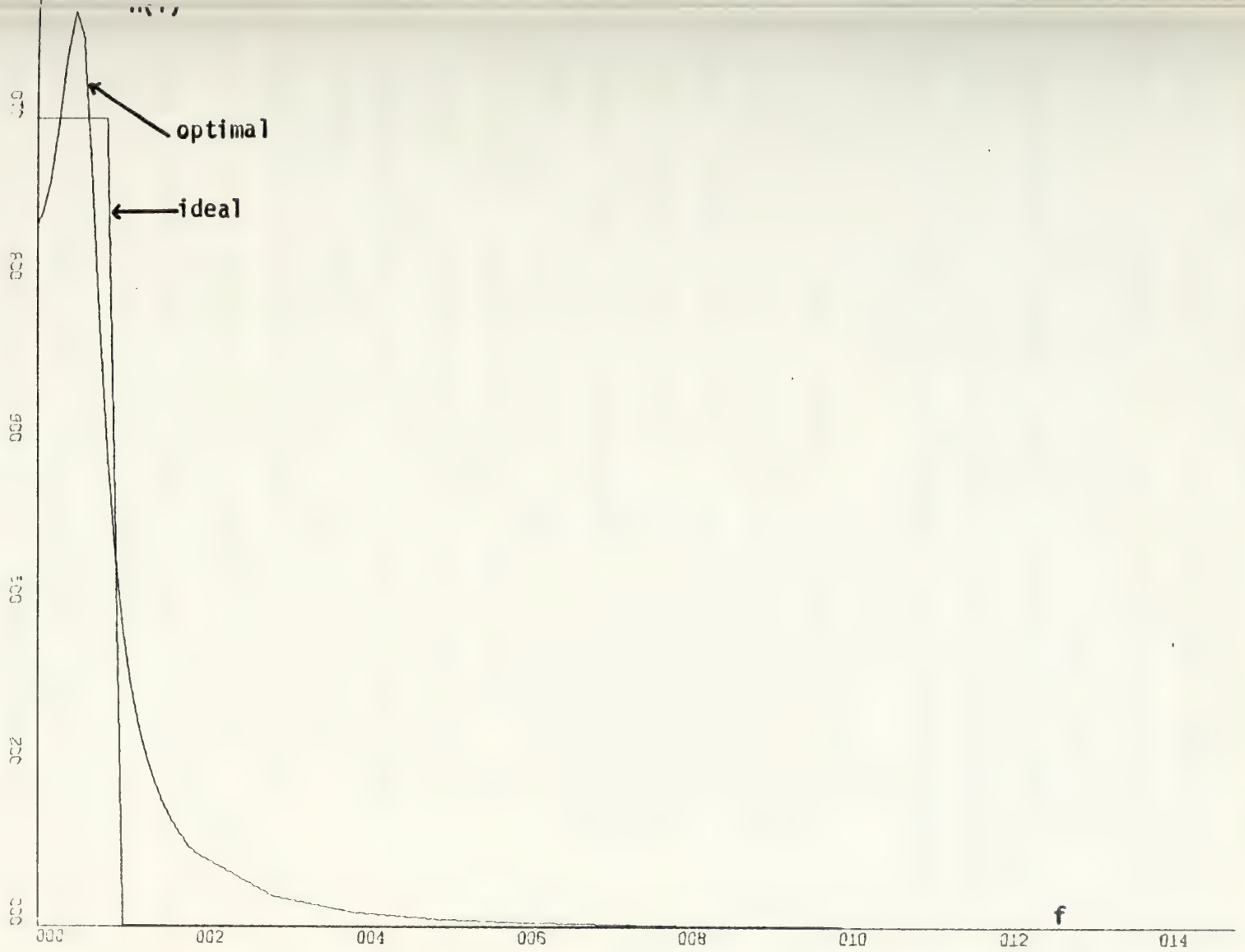


Figure 2.70. First "tooth" of the transfer function of a second order, low pass, sampled analog filter whose feed forward coefficients were held constant at (1,2,1). Notice the zero at half the sampling frequency (010 on the horizontal axis).

decided on two CCD's from the 747 lot, made by TRW. The operating curves which seemed best were curves for the input on gate 3 and making V_{G1} serve as the constant parametric voltage. During the course of the experiments, we constantly monitored the actual output waveform of the CCD since they tended to drift off the proper bias levels, and also for filtering reasons since feedback often causes large signals to appear internally.

B. More hardware was needed. Instead of needing only one sample-and-hold device, we initially used two ... one to restore the analog signal after the first CCD, and one to restore the analog signal after the second CCD. This did not turn out to be adequate, and another sample-and-hold device was inserted at the input in order to be sure that theoretical cancellation would take place. That is, at the summers, we insured that all THREE input signals were of a sampled-and-held form. Also required were additional level shifters. Naturally, the output level had to be shifted so that the ac waveform out had an average value of zero. This was done by using two "up" shifters following each CCD. Also necessary were two "down" level shifters since the input of each CCD had to be shifted substantially further negative in order for the proper bias conditions to exist. These level shifters were built from operational amplifiers with high input impedance. As a matter of fact, we had improved the input impedance to the summers and level shifters so that we did not have to make the "correction

factor" discussed in Section 2.2.1. The only correction we had to make in the setting of the potentiometer coefficients was for the voltage gain of the CCD.

Figure 2.71 shows the physical arrangement of circuit components for the evaluation of the second order sampled analog recursive comb filter.

We planned the experimental work by investigating Figures 2.55 through 2.59 for several transfer functions which would show feasibility of sampled analog recursive filters and also illustrate the power of the b-plane triangle of stability concept. Table 2.5 shows our complete experimental plan. Some points on the b-plane were selected for their monotonic transfer functions, while others were chosen for their large ripple. All the filters tested were low pass, but since the high pass triangle contours are precisely the "upside-down, mirror image" of the low pass triangle contours, we believe that little generality was lost. Some pairs of b's were selected which were close to the boundary of the triangle, and their transfer functions exhibited, in general, higher selectivity (Q), than those chosen nearer the center of the triangle (the origin).

There is one difficulty in setting coefficient values which should be passed on to those who will make further measurements. For that reason, it is included here:

A. Settings of b_1 may range from -2 to +2. One good way of deriving the proper potentiometer value is to

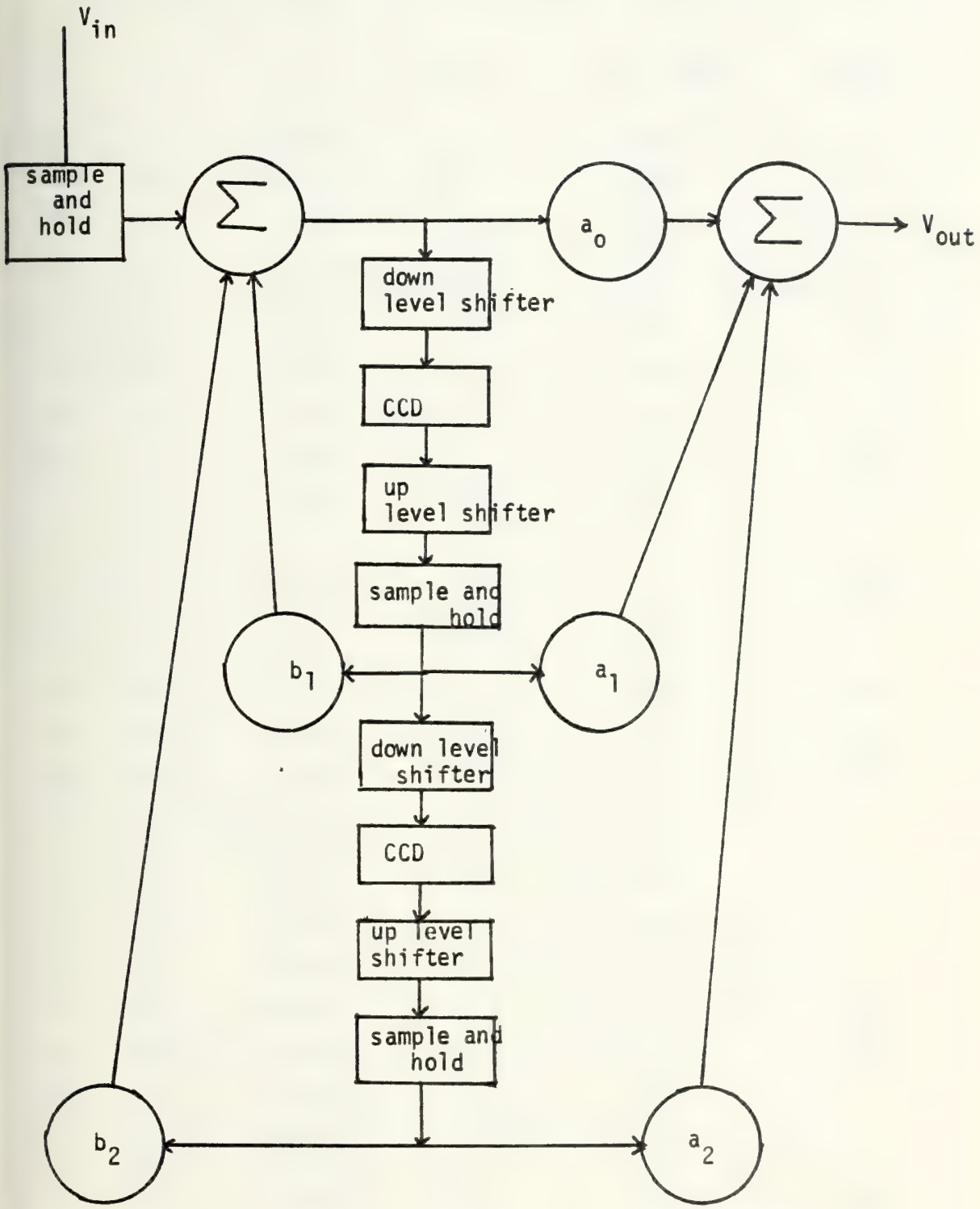


Figure 2.71. Experimental set-up for second order recursive sampled analog filter using CCD's as the delay elements.

DATA TAKING PLAN: $a_o = 1$, $a_1 = 2$, $a_2 = 1$, $b_o = 1$

b_1	b_2	f_c	ripple	Data taken	Figure
0.0	0.0	$0.182f_s$	0 %	YES	
0.0	-0.3	$0.09\ f_s$	0 %	YES	2.73
0.0	-0.7	$0.028f_s$	0 %	YES	2.72
0.0	+0.3	$0.277f_s$	4.6 %	YES	2.74
0.0	+0.7	$0.286f_s$	32.6 %	YES	2.75
-0.3	-0.3	$0.052f_s$	0 %	YES	2.76
-0.3	+0.1	$0.174f_s$	0 %	YES	
-0.3	+0.3	$0.238f_s$	4 %	YES	2.77
-0.3	+0.7	$0.258f_s$	32.4 %	YES	2.78
+0.3	-0.3	$0.124f_s$	0 %	YES	2.79
+0.3	+0.1	$0.260f_s$	0 %	YES	
+0.3	+0.3	$0.312f_s$	4 %	YES	2.80
+0.3	+0.7	$0.313f_s$	32.4 %	YES	2.81
-0.6	-0.3	$0.013f_s$	0 %	NO	
-0.6	+0.1	$0.110f_s$	0 %	YES	2.82
-0.6	+0.3	$0.191f_s$	2 %	YES	
-0.6	+0.7	$0.229f_s$	31.5 %	YES	2.83
+0.6	-0.3	$0.175f_s$	0 %	NO	
+0.6	+0.1	$0.282f_s$	0 %	YES	
+0.6	+0.3	$0.342f_s$	2 %	YES	2.84
+0.6	+0.7	$0.341f_s$	31.5 %	YES	
-1.0	+0.1	$0.019f_s$	0 %	NO	
-1.0	+0.3	$0.091f_s$	0 %	YES	2.85
-1.0	+0.7	$0.184f_s$	29 %	YES	2.86
+1.0	+0.1	$0.285f_s$	0 %	NO	
+1.0	+0.3	$0.367f_s$	0 %	YES	
+1.0	+0.7	$0.379f_s$	29 %	YES	2.87
-1.3	+0.45	$0.059f_s$	0 %	YES	2.88
-1.3	+0.8	$0.148f_s$	34 %	YES	
+1.3	+0.45	$0.410f_s$	0 %	YES	
+1.3	+0.8	$0.401f_s$	34 %	YES	2.89
-1.6	+0.75	$0.085f_s$	17 %	YES	
+1.6	+0.75	$0.449f_s$	17 %	YES	

TABLE 2.5 Data Taking Plane

calculate the recursive equations for each output in time as a function of an input unit pulse or unit step. Then, while observing the output waveform on the oscilloscope, impress either the unit pulse or the unit step at the input. The unit step input was found to be more convenient. Since a first order filter is simpler, b_1 may be set while a_2 and b_2 are set at zero. This procedure was followed until the magnitude of b_1 exceeded unity. At this point, a first order filter can no longer be designed with such feedback, and the other feedback coefficient, b_2 must be introduced into the design.

The recursive equations we used for setting coefficients whose magnitude exceeded unity are shown in Table 2.6. The experimental procedure is to impress the unit step at the input, observe the output, and then change the potentiometers until the proper output is reached. Record the settings, and proceed with the experimental work.

Problems were found in the cases where the pair (b_2,b_1) were close to the boundary of the triangle. At times, no data at all could be taken since the filter was unstable. Four cases where no data at all was taken are shown in Table 2.5.

Figures 2.72 through 2.89 are tabulated here as representative of the data taken during the course of this experiment. All the graphs are different from their counterparts in the first order case because of the change in the y-scale. It is logarithmic. The "wiggles" in many of the curves occur

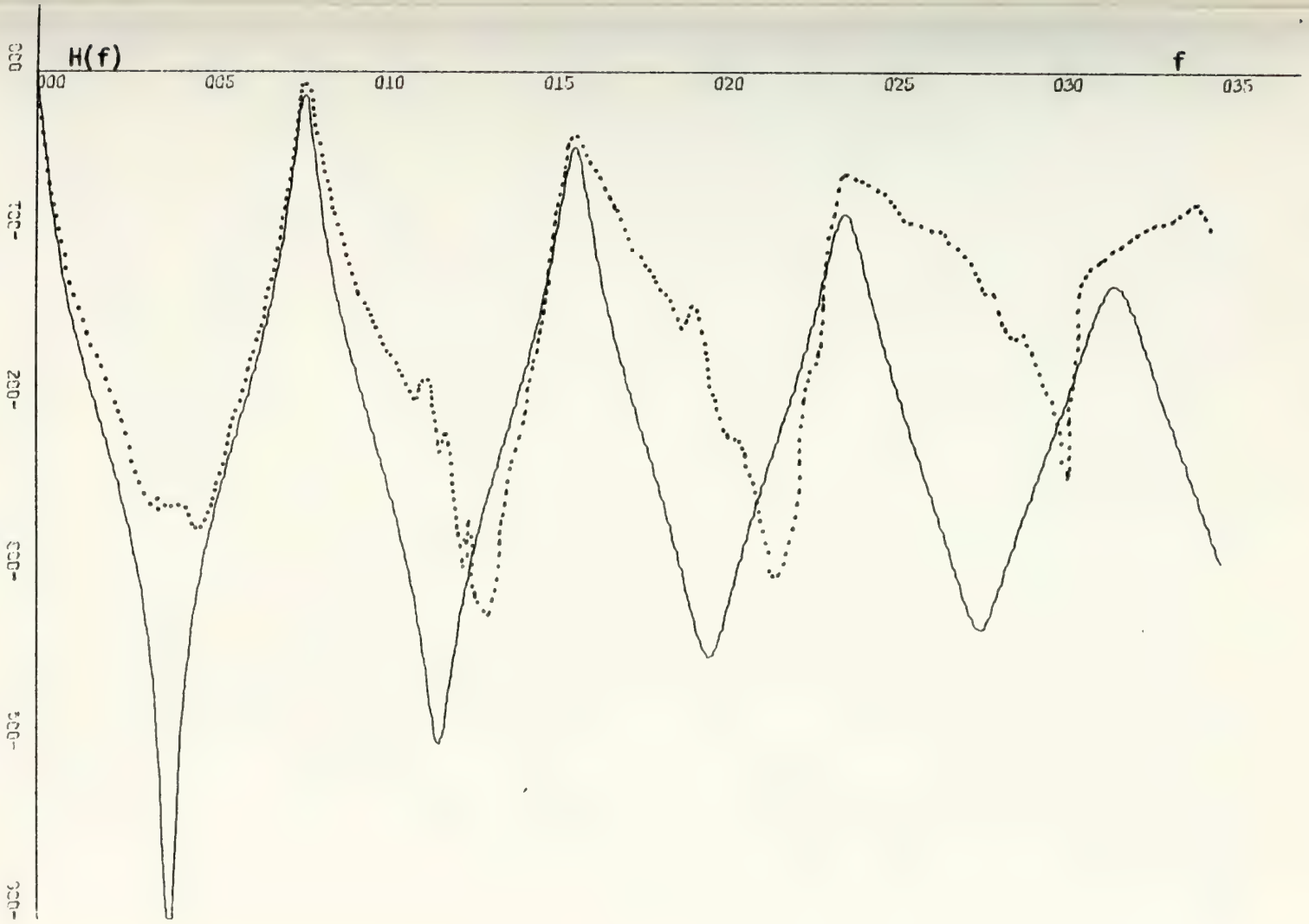


Figure 2.72. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = 0$, $b_2 = -0.7$. Dotted lines are experimental points. (LOG SCALE)

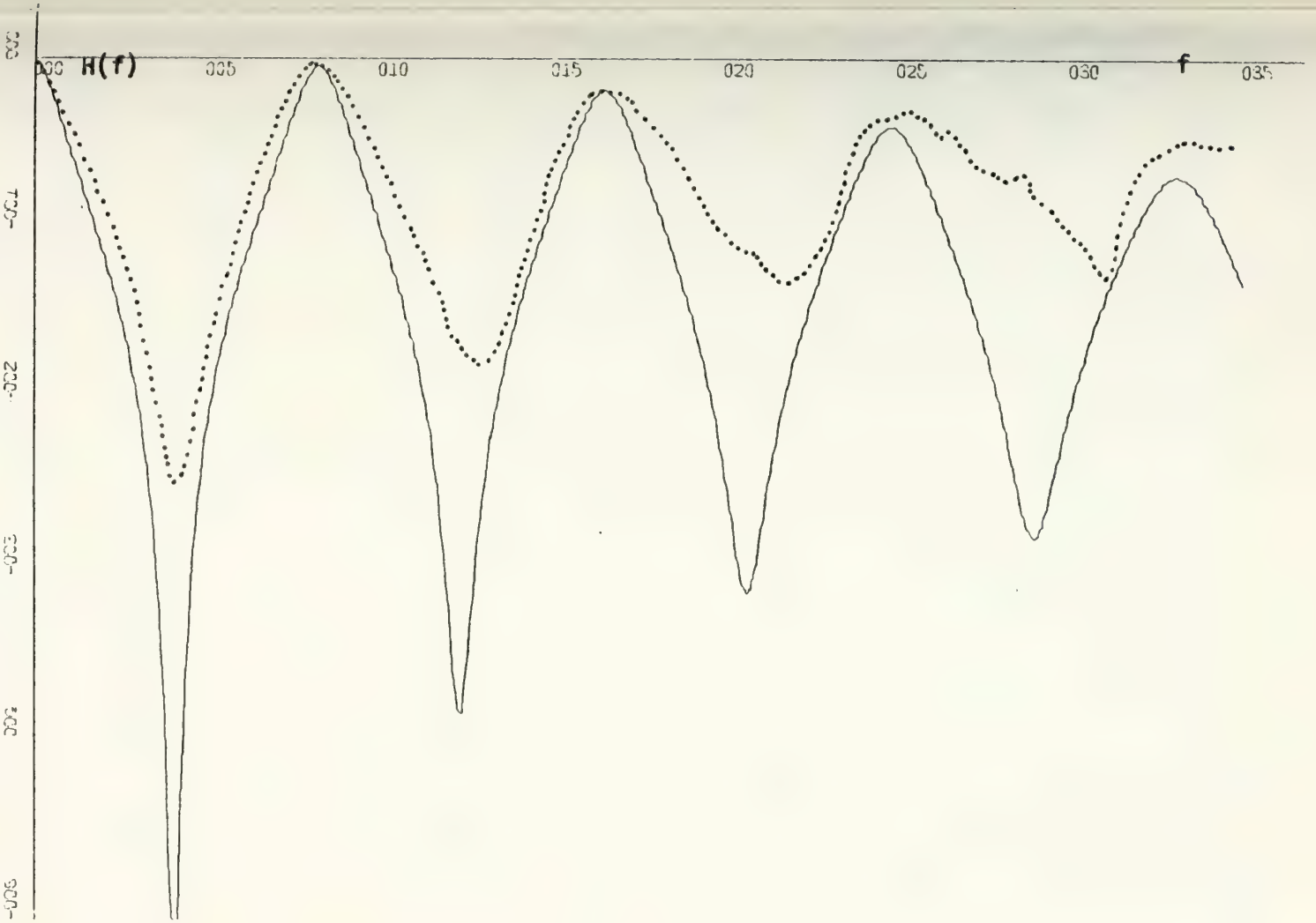


Figure 2.73. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = 0$, $b_2 = -0.3$ Dotted lines are experimental points. (LOG SCALE)

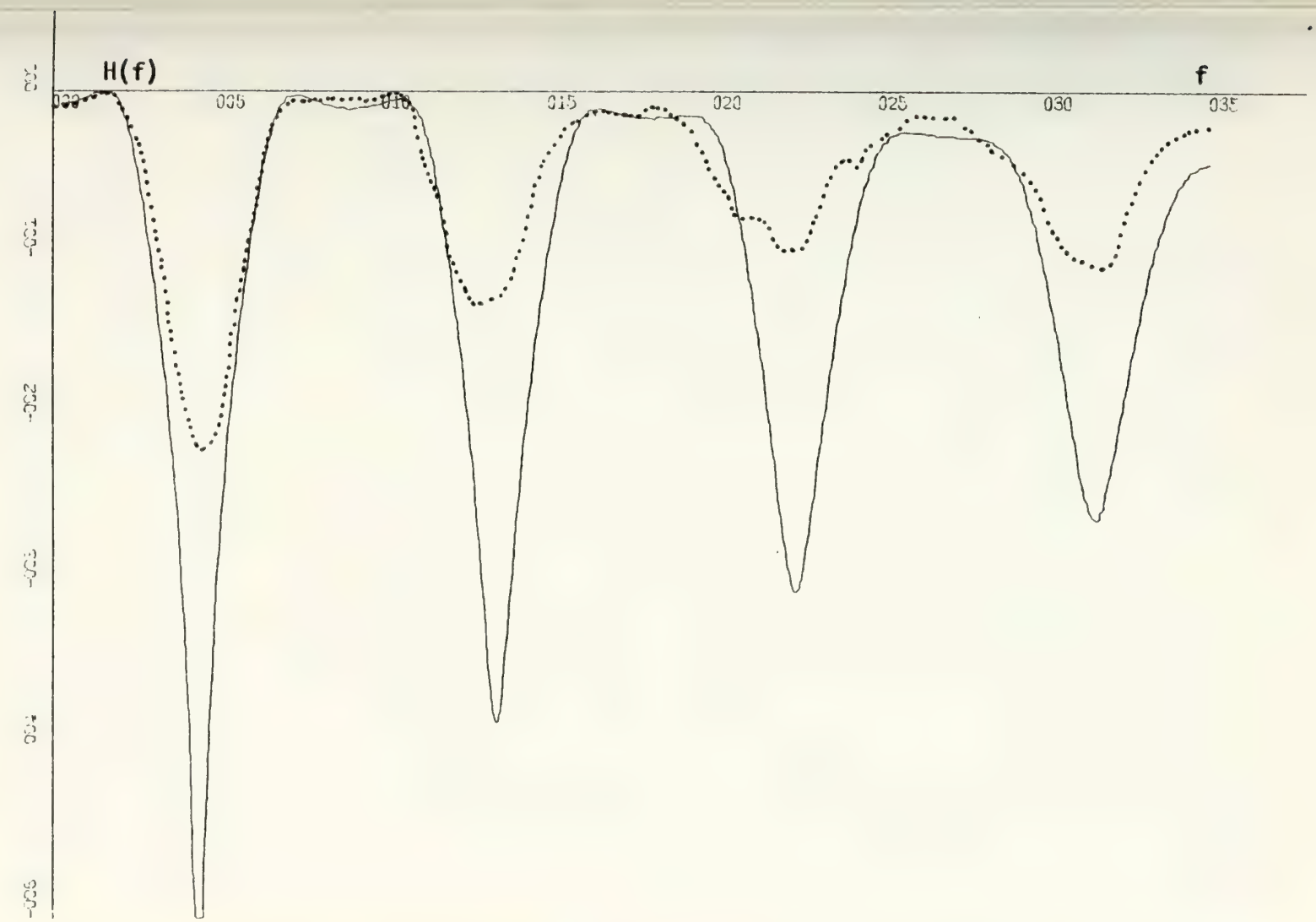


Figure 2.74. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = 0$, $b_2 = 0.3$. Dotted lines are experimental points. (LOG SCALE)

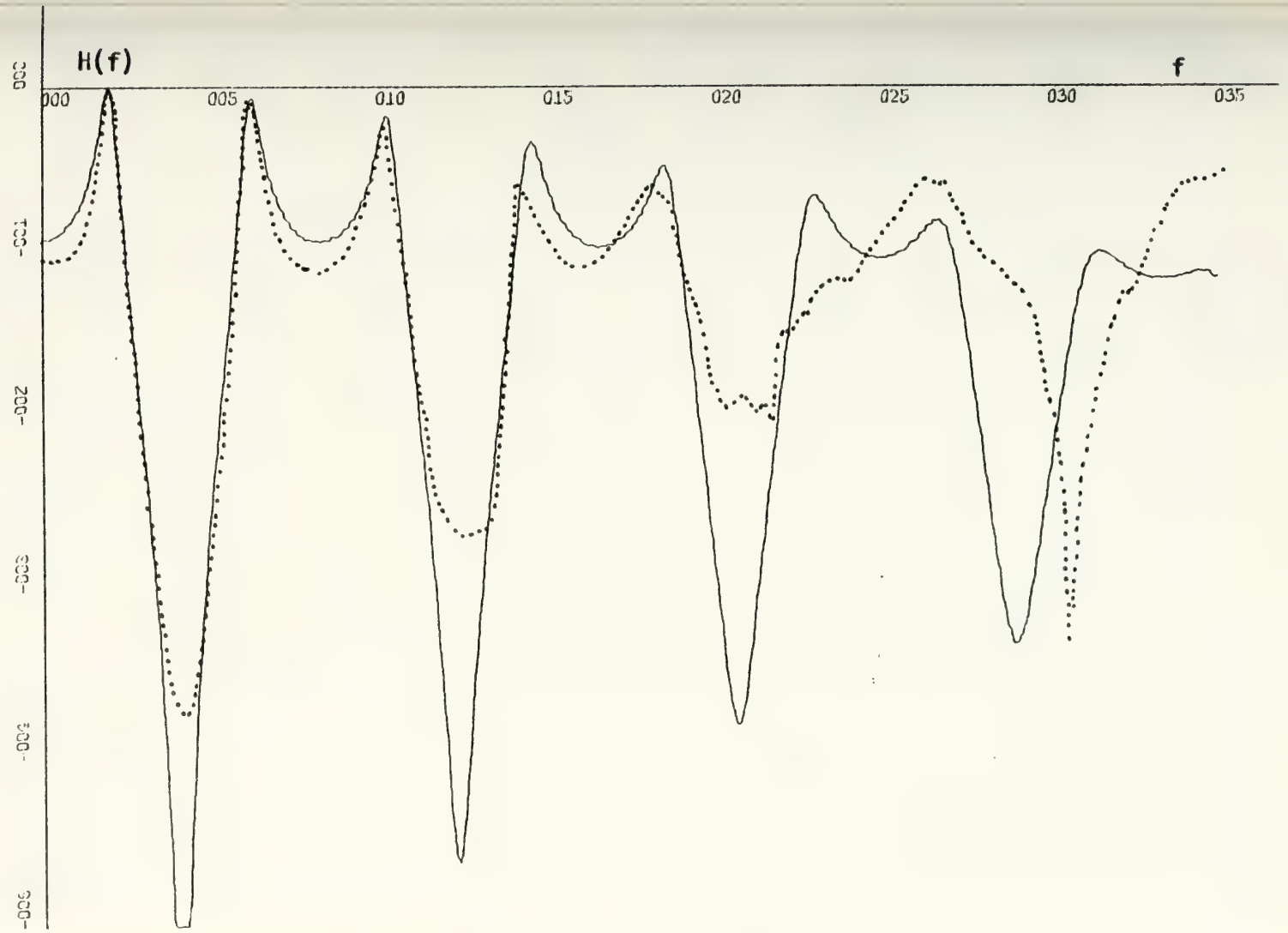


Figure 2.75. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = 0$, $b_2 = 0.7$. Dotted lines are experimental points. (LOG SCALE)

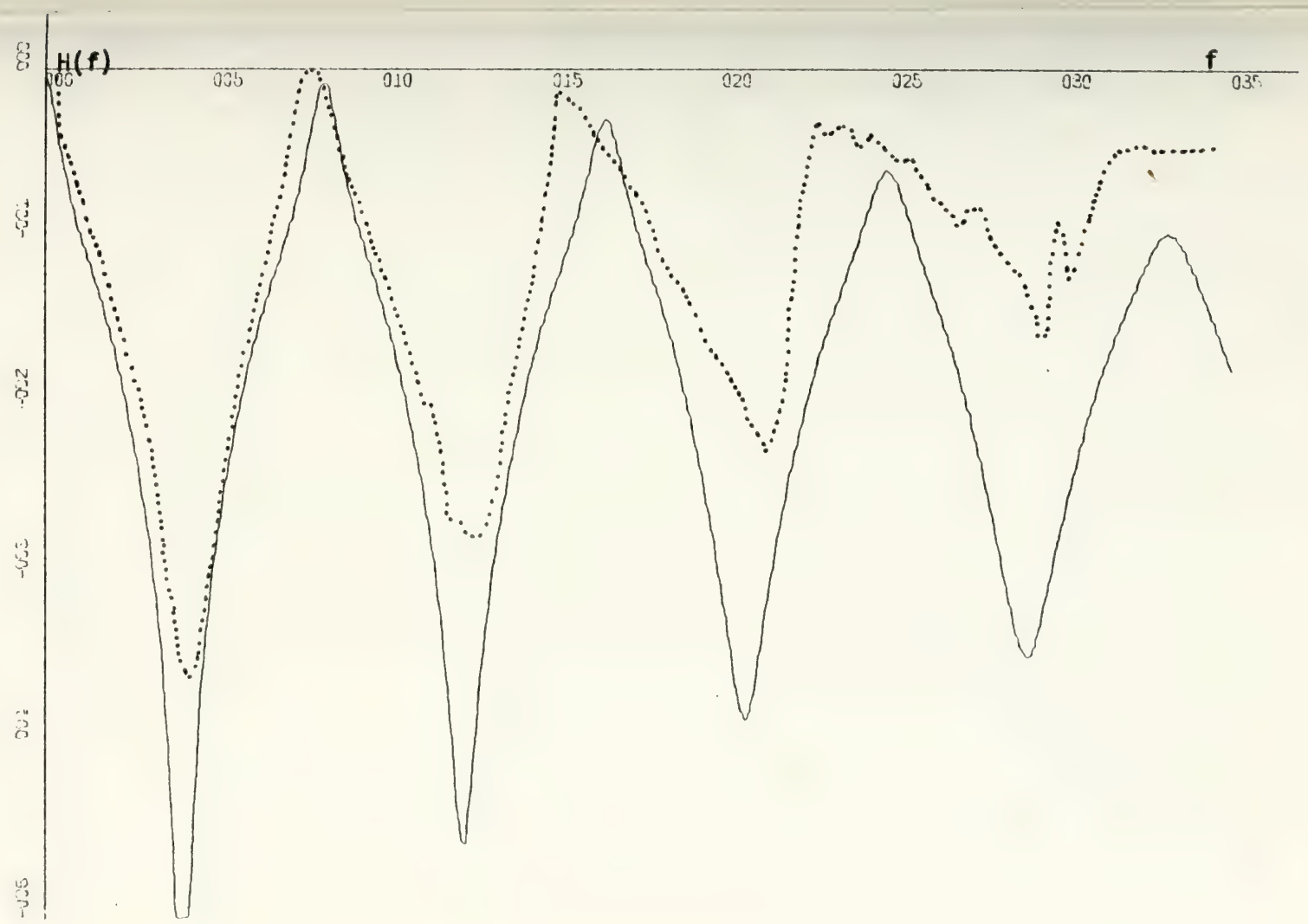


Figure 2.76. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -0.3$, $b_2 = -0.3$. Dotted lines are experimental points. (LOG SCALE)

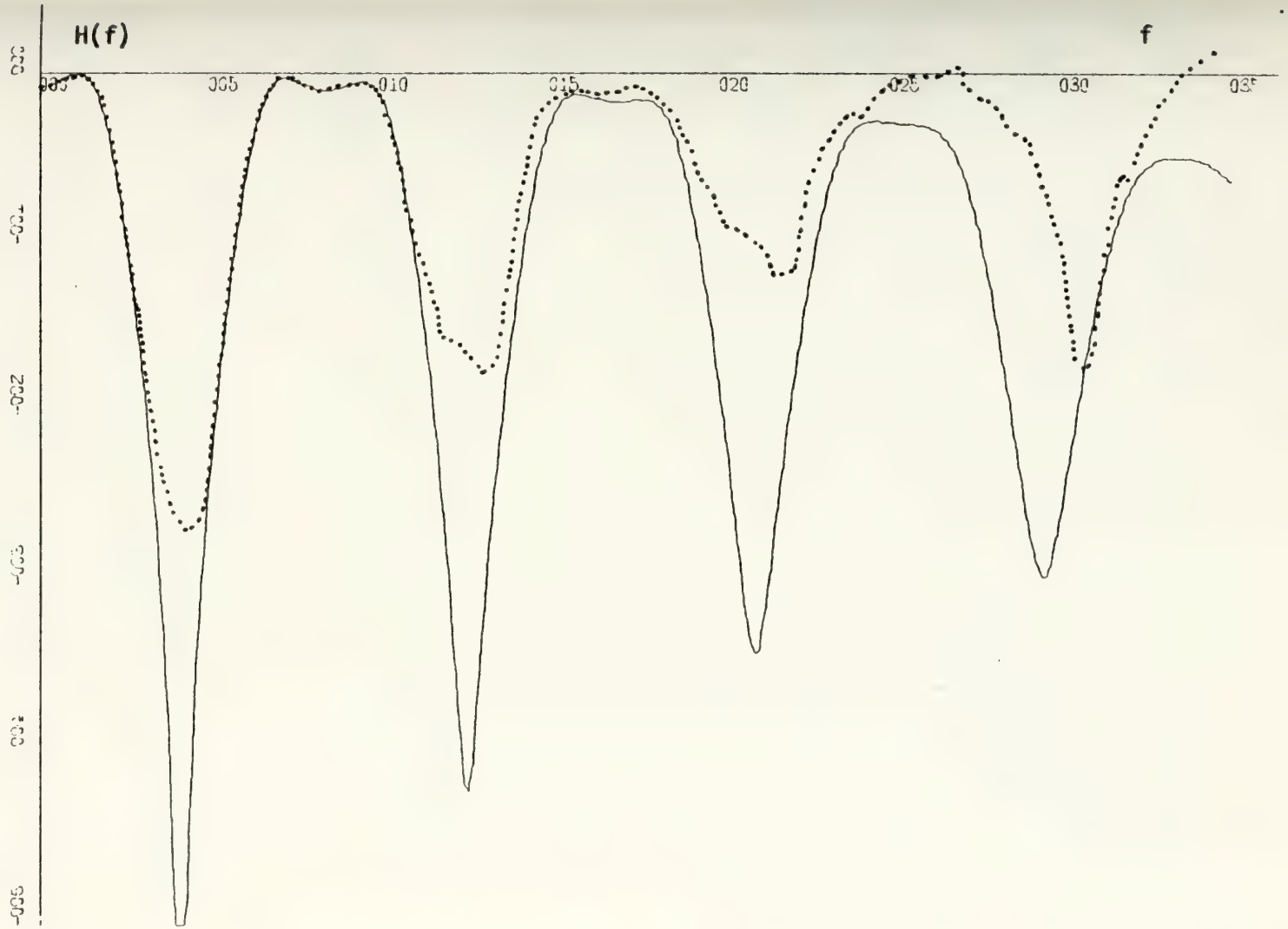


Figure 2.77. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -0.3$, $b_2 = +0.3$. Dotted lines are experimental points. (LOG SCALE)

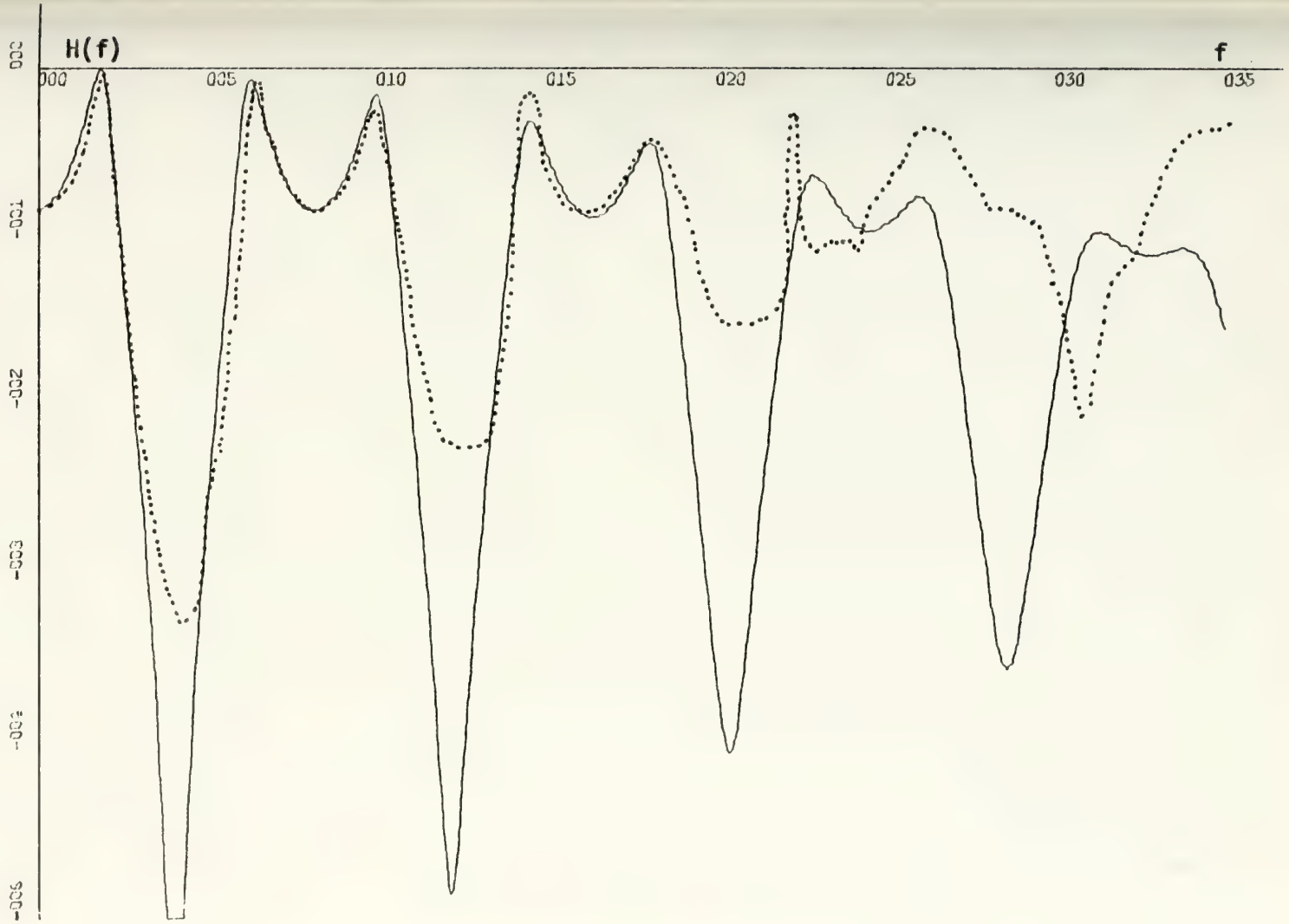


Figure 2.78. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -0.3$, $b_2 = +0.7$. Dotted lines are experimental points. (LOG SCALE)

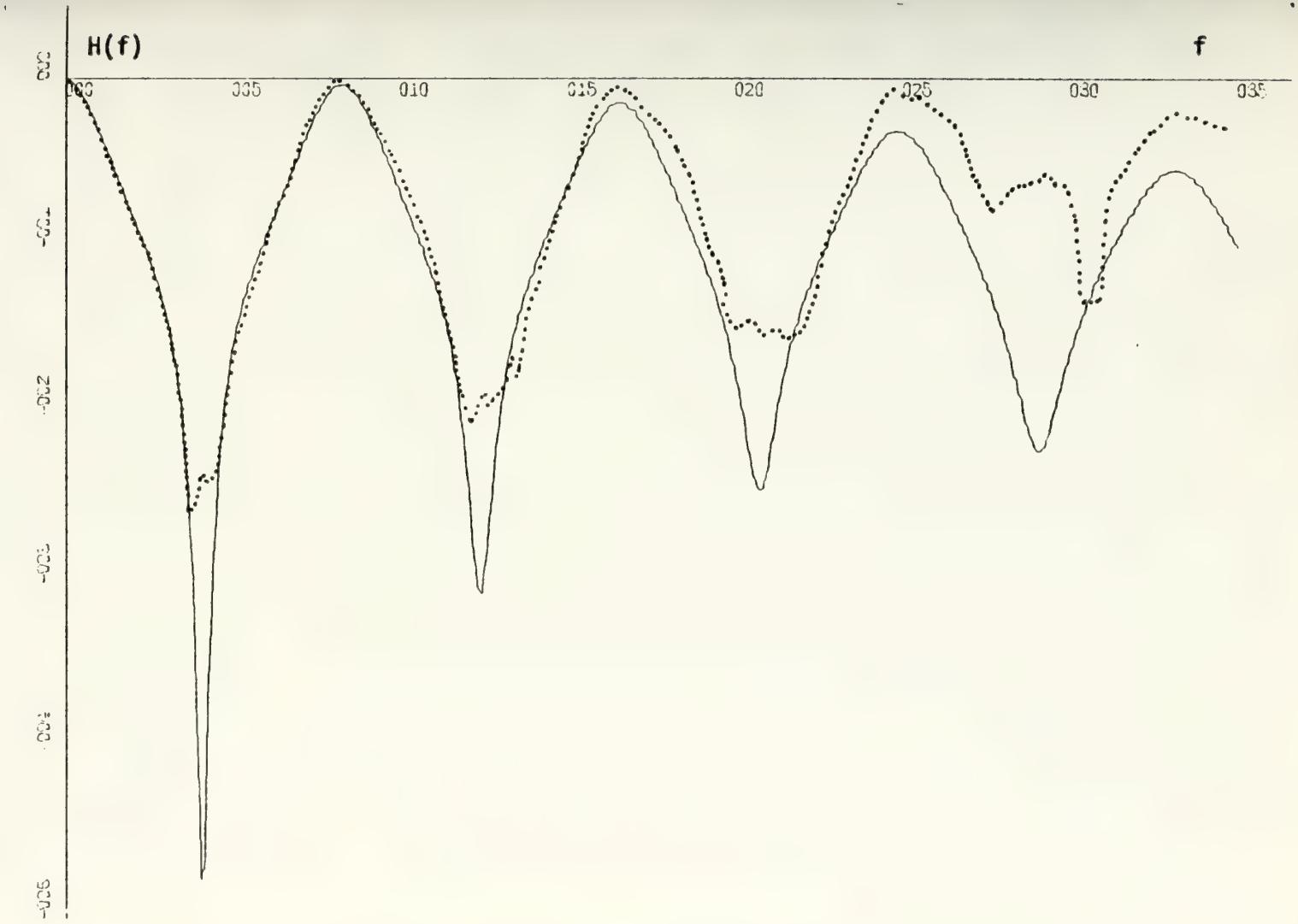


Figure 2.79. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = +0.3$, $b_2 = -0.3$. Dotted lines are experimental points. (LOG SCALE)

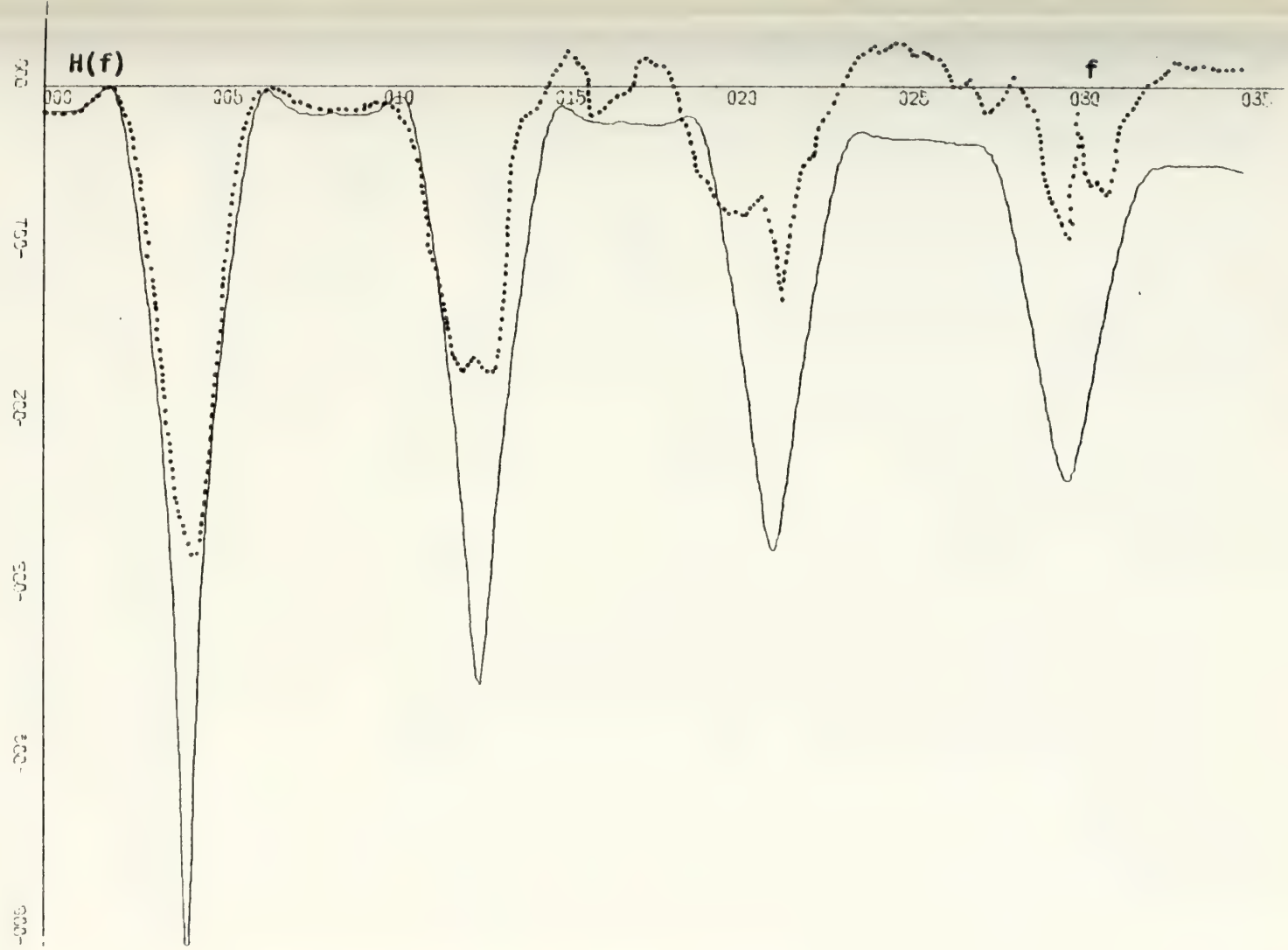


Figure 2.80. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = +0.3$, $b_2 = +0.3$. Dotted lines are experimental points. (LOG SCALE)

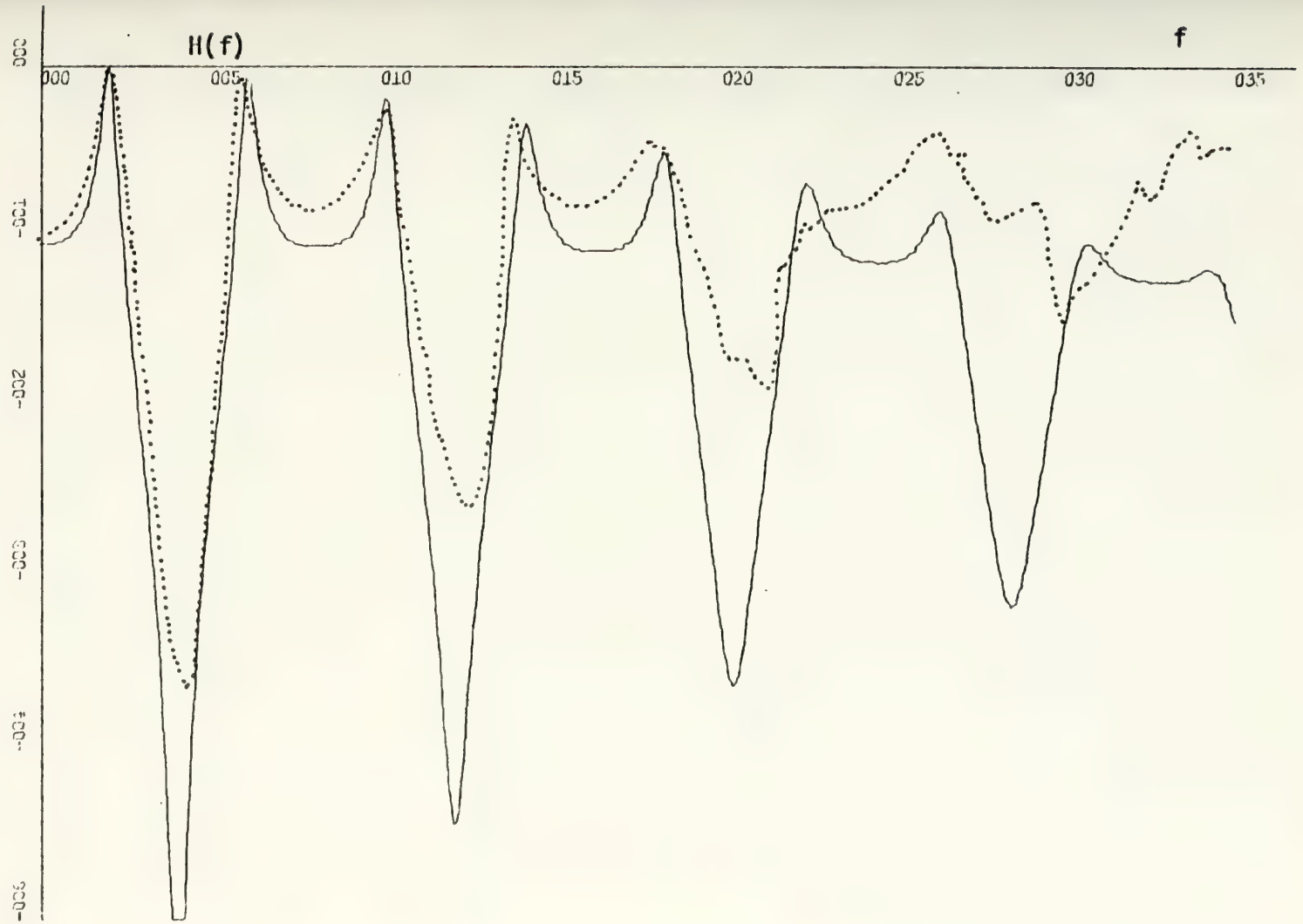


Figure 2.81. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = +0.3$, $b_2 = +0.7$. Dotted lines are experimental points. (LOG SCALE)

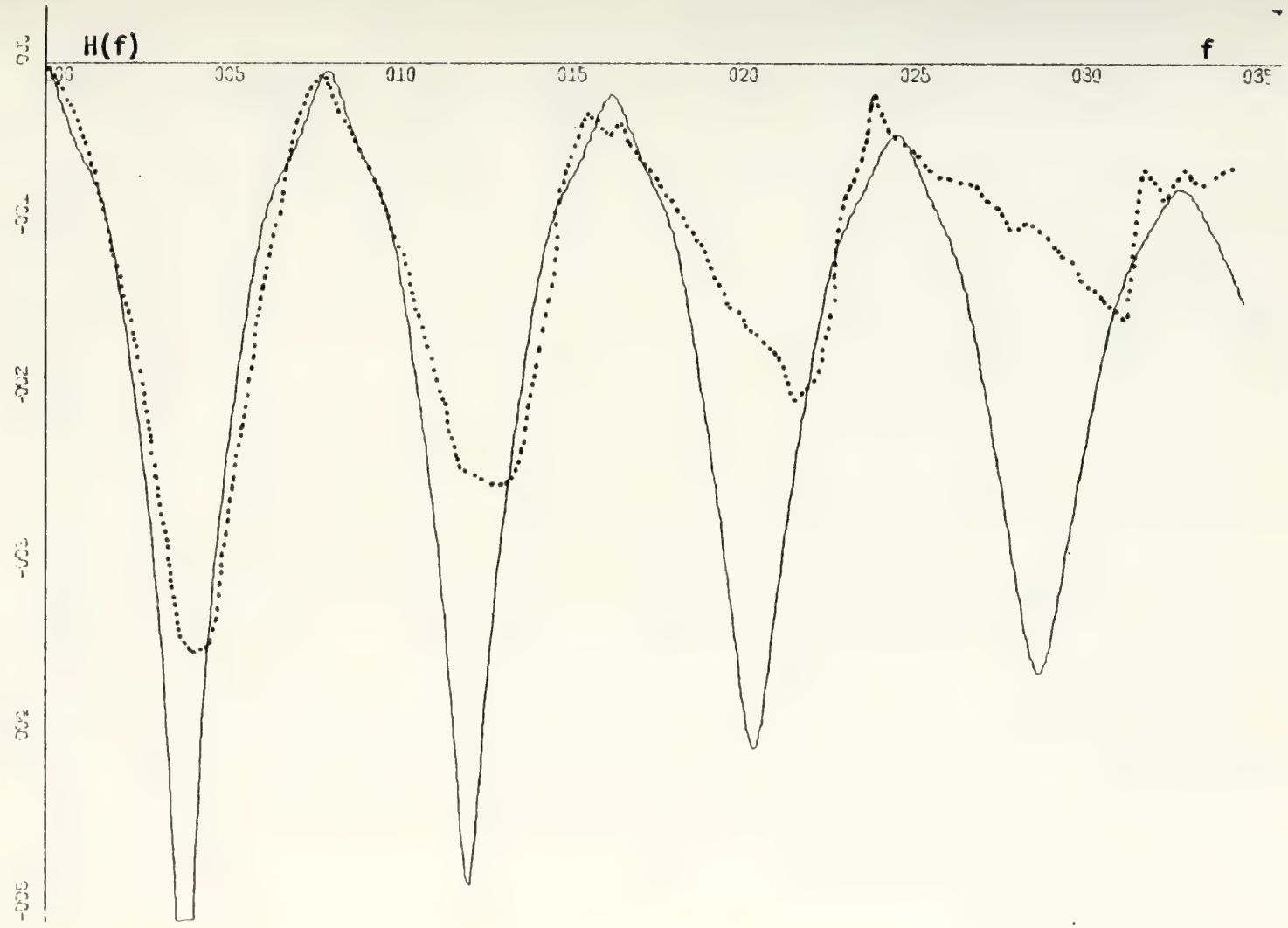


Figure 2.82. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -0.6$, $b_2 = +0.1$. Dotted lines are experimental points. (LOG SCALE)

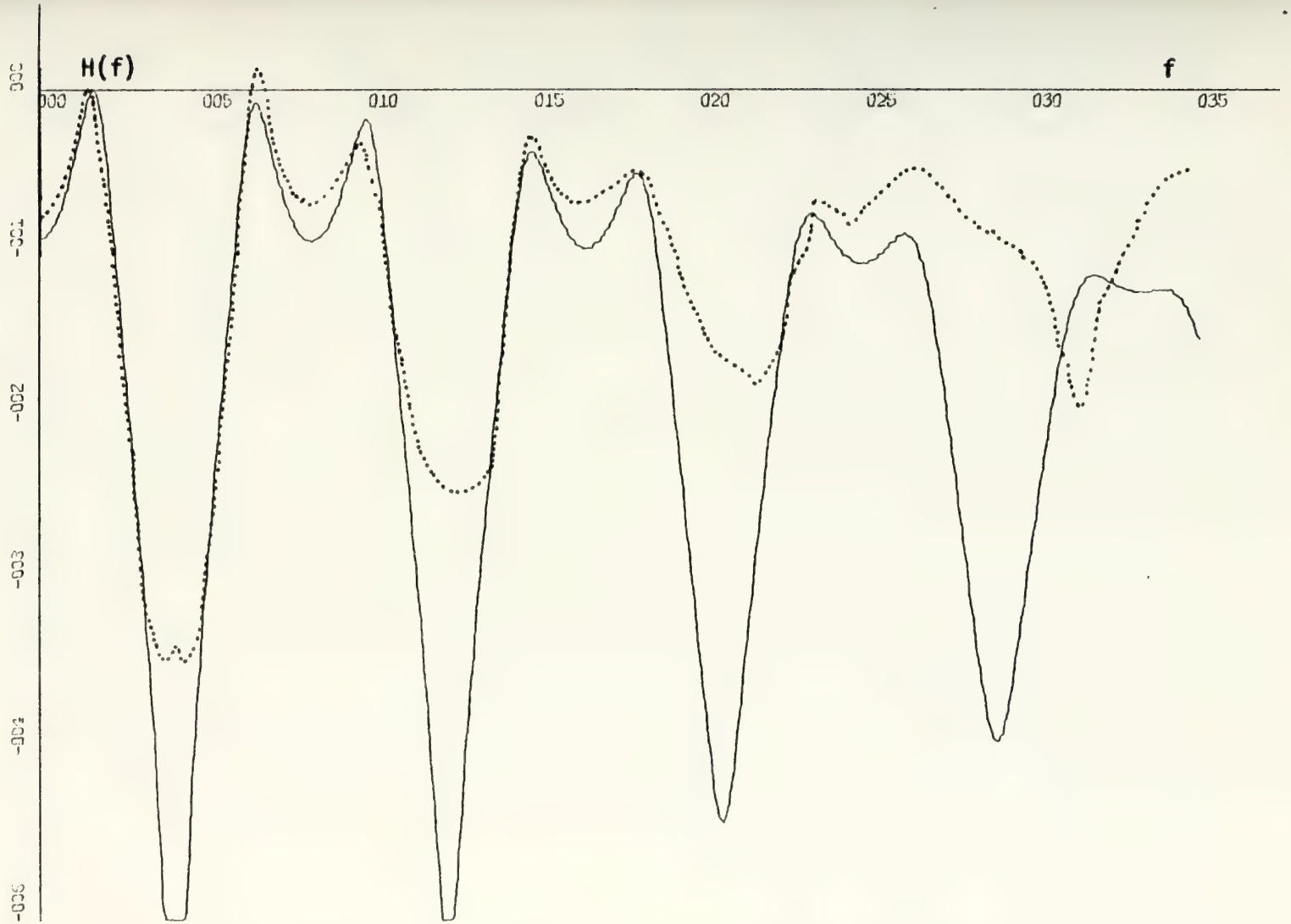


Figure 2.83. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -0.6$, $b_2 = 0.7$. Dotted lines are experimental points. (LOG SCALE)

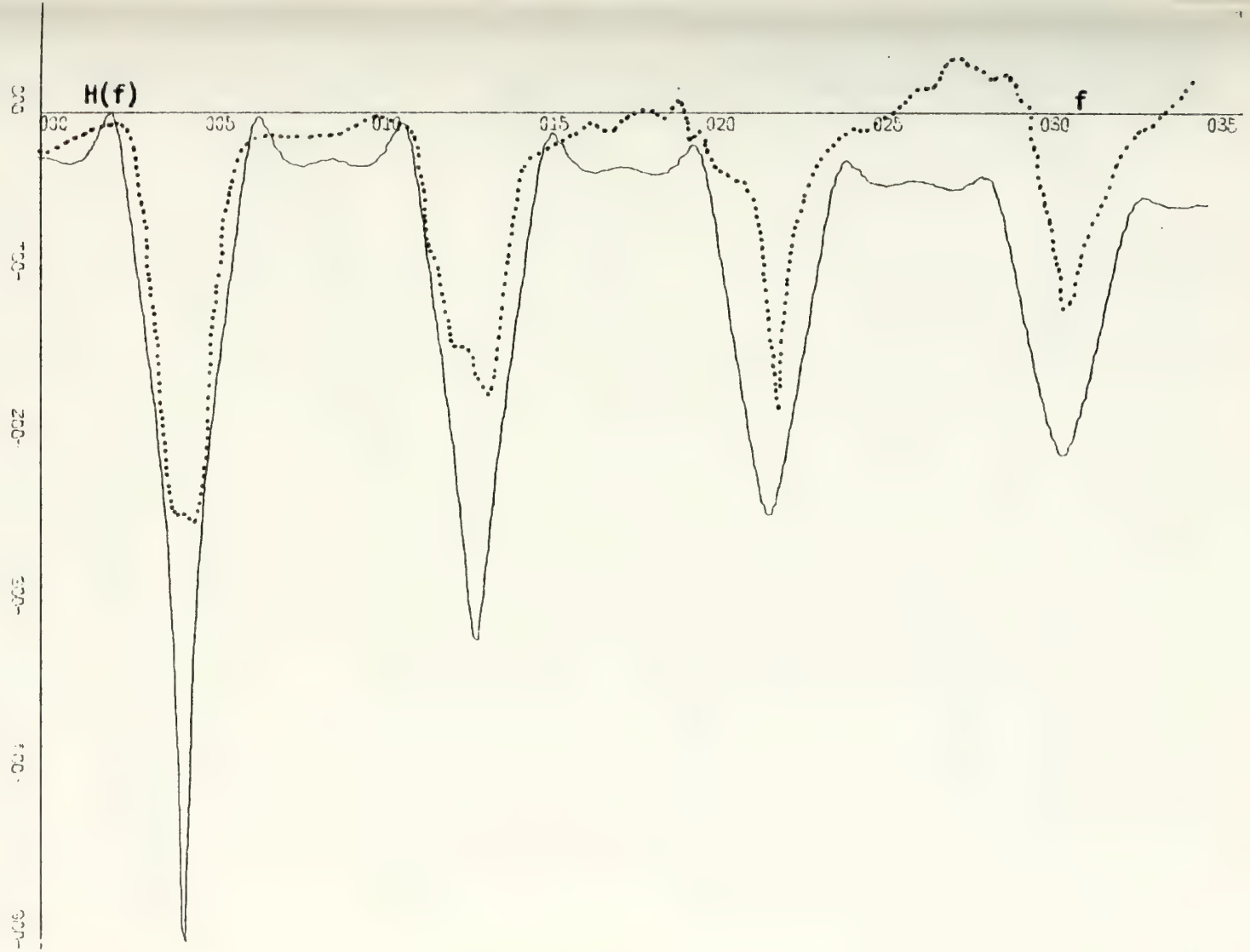


Figure 2.84. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = +0.6$, $b_2 = +0.3$. Dotted lines are experimental points. (LOG SCALE)

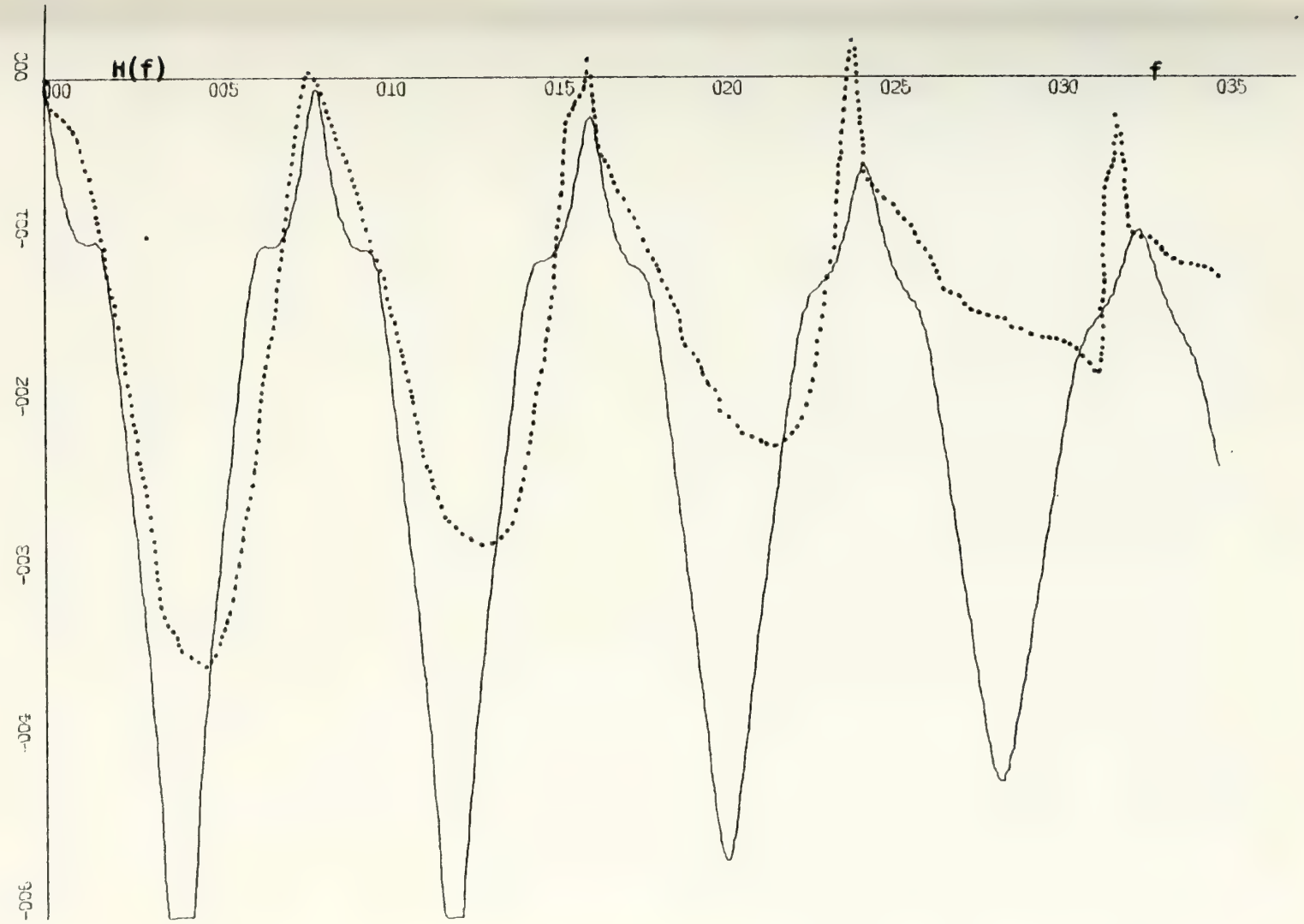


Figure 2.85. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -1.0$, $b_2 = +0.3$. Dotted lines are experimental points. (LOG SCALE)

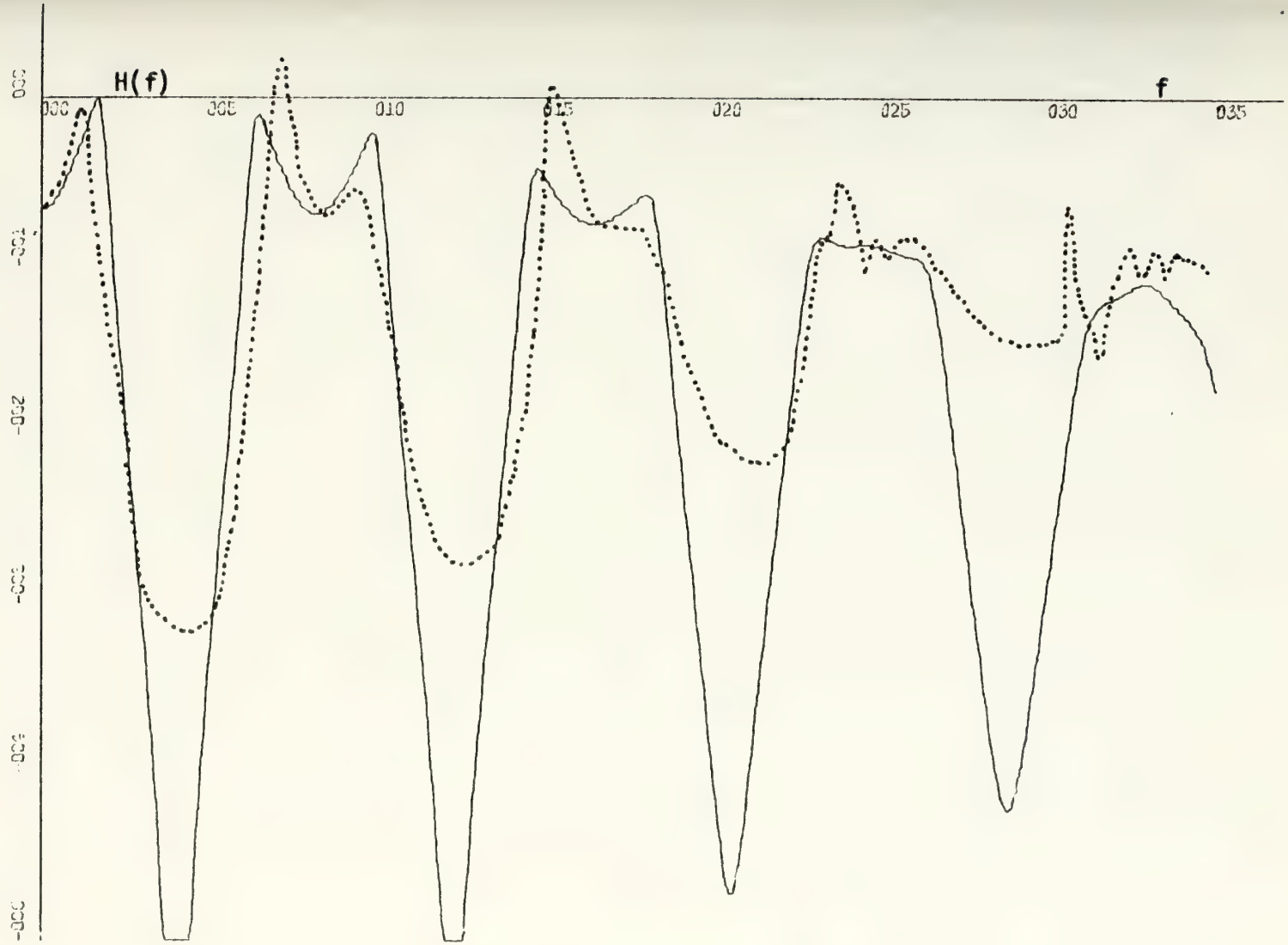


Figure 2.86. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -1.0$, $b_2 = +0.7$. Dotted lines are experimental points. (LOG SCALE)

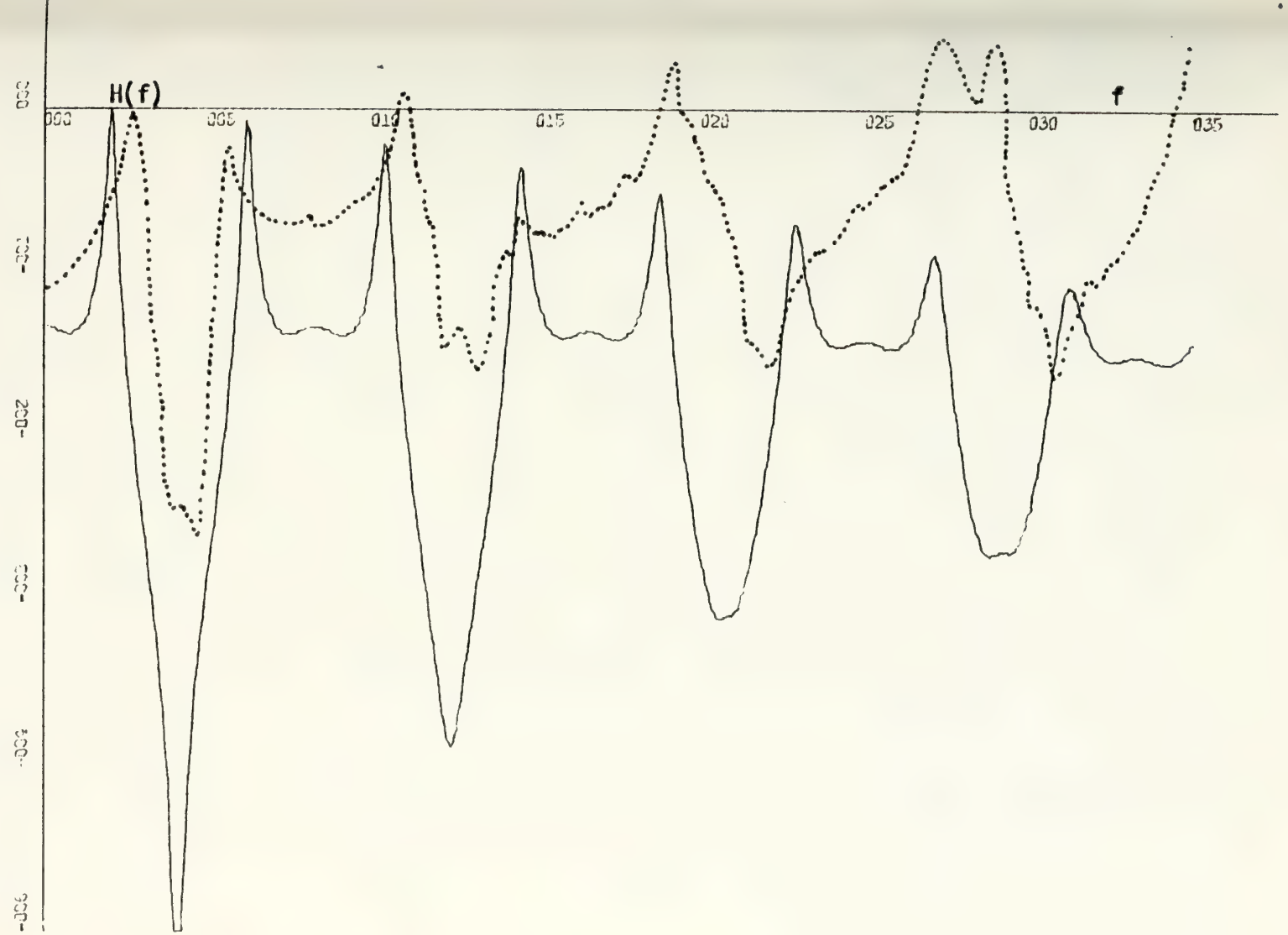


Figure 2.87. Theoretical and Experimental transfer function for a second order recursive comb filter, $b_1 = +1.0$, $b_2 = +0.7$. Dotted lines are experimental points. (LOG SCALE)

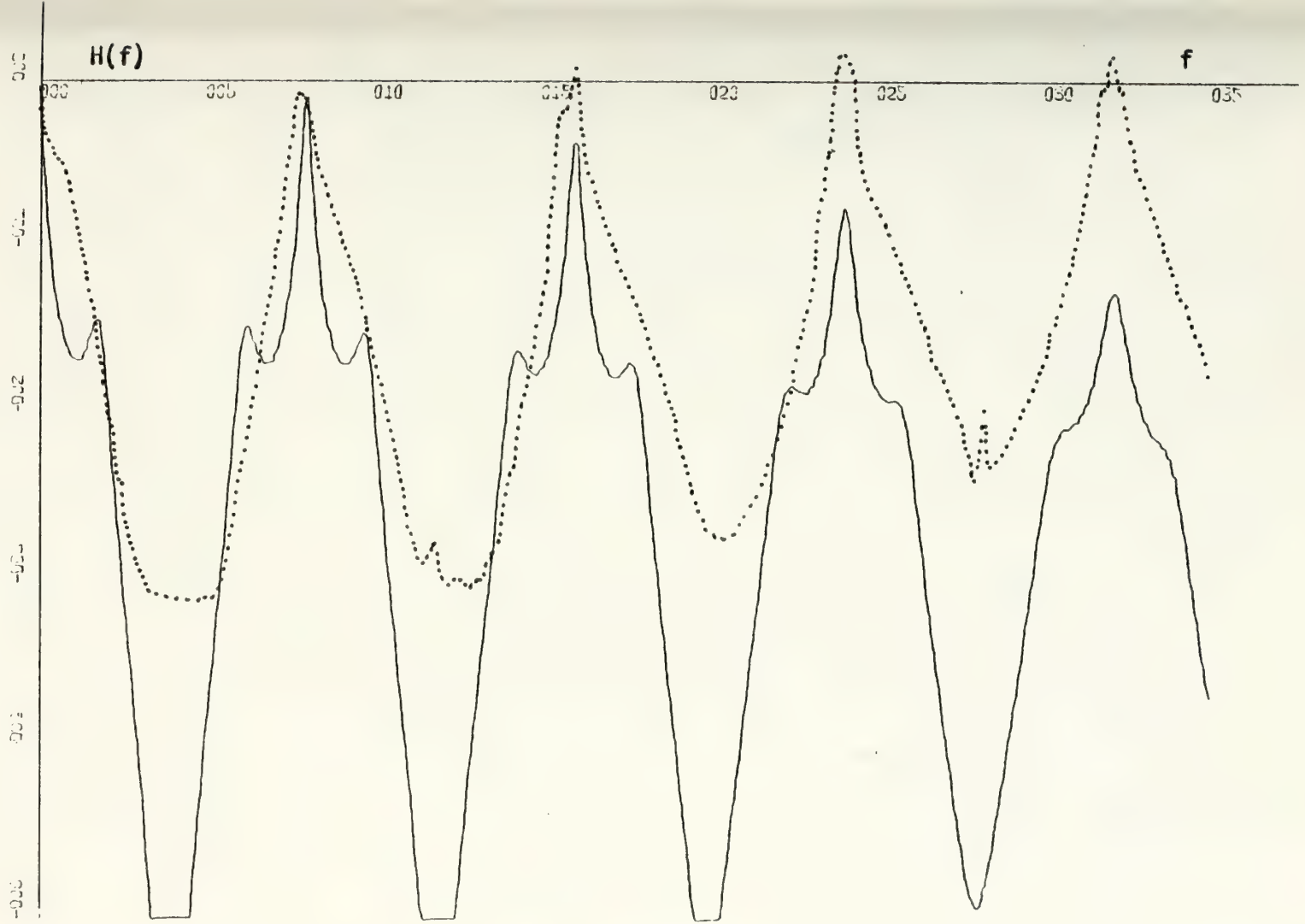


Figure 2.88. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = -1.3$, $b_2 = 0.45$. Dotted lines are experimental points. (LOG SCALE)

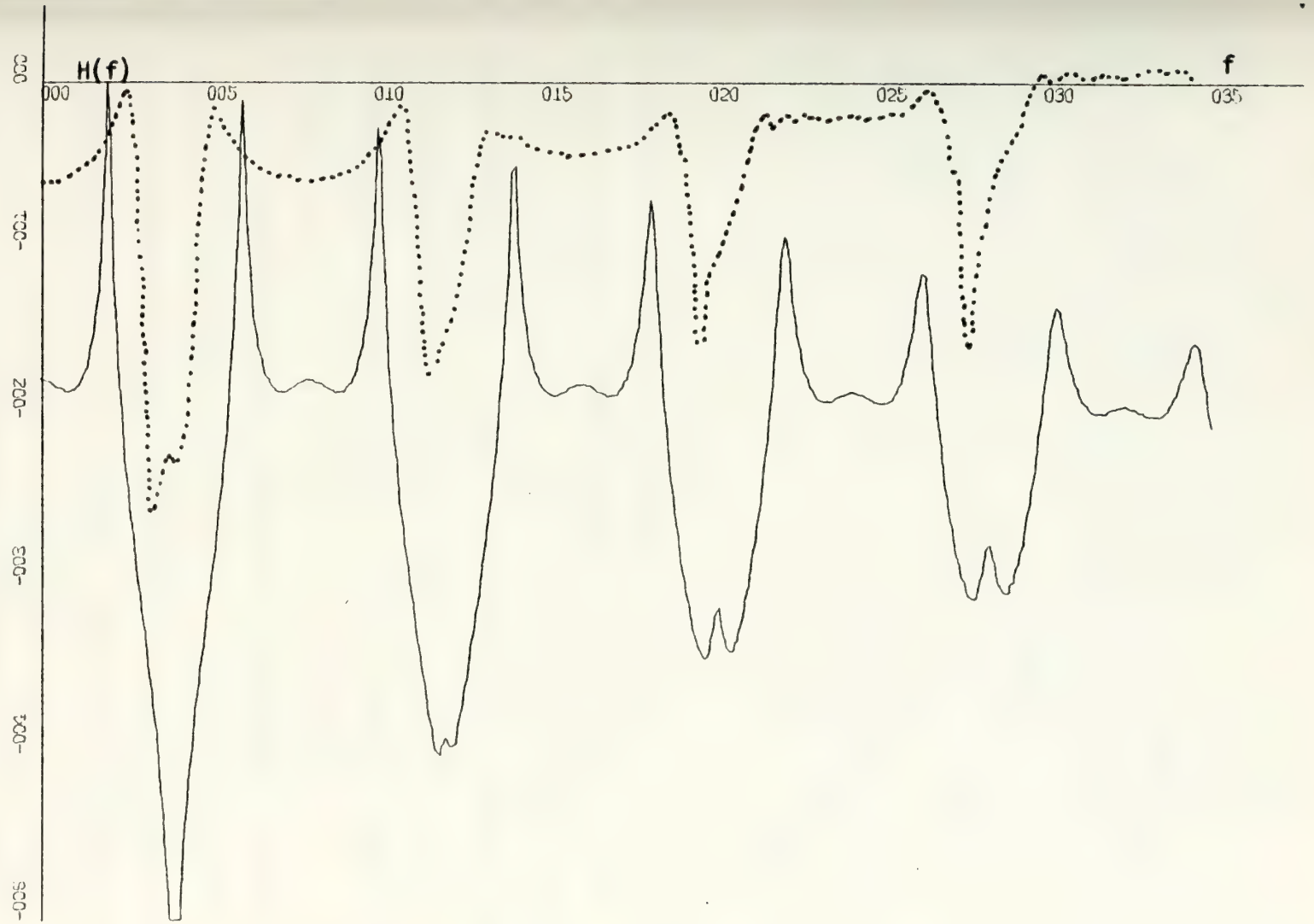


Figure 2.89. Theoretical and experimental transfer function for a second order recursive comb filter, $b_1 = +1.3$, $b_2 = +0.8$. Dotted lines are experimental points. (LOG SCALE)

TABLE 2.6

Recursive equations used to set b_1 when $b_1 = 1$.

$a_o = 1,\ a_1 = 0,\ a_2 = 0,\ b_o = 1,\ b_1 = 1,\ b_2 = 1$

TIME	INPUT	OUTPUT
0	1	1
1	1	$1 + b_1$
2	1	$(1+b_1)b_1+1+b_2 = 1+b_1+b_1^2+b_2$
3	1	$b_1+b_1^2+b_1^3+2b_1b_2+b_2$

at a low response level and are the results of the system noise as well as imperfections in the filter itself. This was not apparent in the first order experiments because of the linear scale. In many cases agreement between theory and experiment was extremely good ... notwithstanding the somewhat erratic results seen at low output level and at frequencies near the Nyquist limit (10 kHz).

All the theoretical curves are corrected for the sample-and-hold effect and also for the inefficiency of the device which we conservatively estimated at 0.001. Thus, the theoretical curves are somewhat different from what one might expect. The inefficiency causes the null (at $f_s/16$, $3f_s/16$, $5f_s/16$, ...) to be progressively poorer as frequency is increases. It also causes the peaks of the output to change with frequency and sometimes, we get a triple peak as in Figure 2.88. The triple peak would not be possible if

inefficiency were zero since, in that case, the values of the feedforward coefficients would always be precisely (1,2,1) so that both zeros in the transfer function would lie at z = -1 on the z-plane. But the effect of inefficiency moves the zeros away from that point, thus causing multiple peaks in the passband and a deterioration of the null in the stopband.

Of course, the sample-and-hold effect is a familiar sin x/ x multiplicative effect.

Once again, as we did in the experiments in first order, we find that several transfer functions have the disturbing tendency to rise with increasing frequency. We believe that the major factor contributing to this rise in the increase in reactance causing a decrease in b_1 and b_2 at higher frequencies. A quick look at the individual experiments which manifested this phenomenon confirms that the rising occurred mainly when the feedback was positive:

b_1	b_2	rising effect
-0.3	+0.3	mild
+0.3	+0.3	strong
+0.6	+0.3	medium
-1.0	+0.3	medium
+1.0	+0.7	strong
-1.3	+0.45	strong
+1.3	+0.8	strong

Since these same potentiometers were used in previous experiments as well, it is appropriate to determine whether or not

others have found this phenomenon. In a thesis completed here in June, 1975, Ensign B. Freund determined that the RETICON (SAD) serial analog delay device exhibited similar properties when used in first order filters. One common set of equipments used in the different experiments were the potentiometers.

In order to investigate this increase further, we backed off from our assumption of an equivalent second order system and took the more conservative view that our device is a sixteenth order filter. Further, we used a set of coefficients (b_2,b_1) which, when 7.8 kHz was reached (at the input), caused the filter to become unstable. Now this behavior cannot be predicted by the second order translation, but it can be easily understood by a z-plane plot of the sixteen poles, which is shown in Figure 2.90.

In order to determine that this was the true situation, we advanced the frequency to about 9.9 kHz where once again the filter exhibited instability. This instability persisted until past 10.0 kHz and finally subsided at 10.15 kHz. Increasing the frequency further (now beyond the Nyquist rate), we found the filter to be stable until 12.2 kHz was reached. Here once again we observed unstable operation. Our sixteenth order theory told us that we should be prepared for an instability at a frequency which corresponded to the complex conjugate of the instability at 7.8 kHz. The frequency at the complex conjugate is 12.2 kHz.

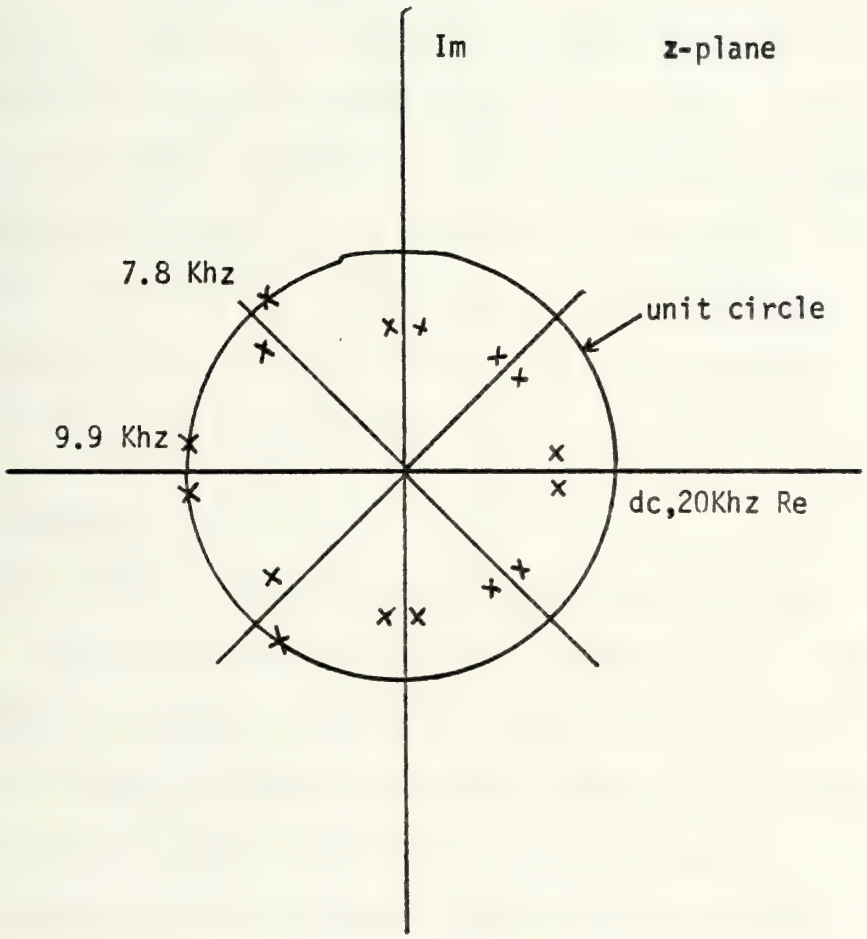


Figure 2.90. The pole plot for our sixteenth order filter which displays instabilities at 7.8 Khz and 9.9 Khz.

Going further, we increased the frequency to 20 kHz, but no other unstable conditions appeared. It was postulated that none would appear until we had traveled all the way around the unit circle (on the z-plane) to the frequency corresponding to 7.8 kHz + one revolution or 27.8 kHz. We found this unstable condition at 28.0 kHz. Thus, we found that our filter, although not conforming to second order translation theory, did conform to sixteenth order filter theory. Of course, there was no way to measure the radial location of the pole at the point corresponding to 7.8 kHz, since the pole may lie outside the unit circle.

SUMMARY

Out of 33 planned experiments, we were able to complete 29. The other four proved so unstable that no data could be taken. Of the 29 sets of data, we find that in many cases, close agreement between theory and experiment existed at least for the first two "teeth." Also, we found that the shapes of the transfer functions were very close (once again, for the first two teeth) to that predicted for them by inspecting the b-plane triangle of stability data. Measurements of cutoff frequency and ripple, especially, were compared to their predicted values with close agreement in most cases. These measurements were made on the first "tooth" of the transfer function which is, of course, the "best behaved" region in the frequency domain. No attempt was made to deal with the standard z or bilinear z transforms.

The standard z transform, once again, produces (for second order) a transfer function which has zeros only at infinity or zero. The bilinear z transform does produce, in general, two zeros and this was, in fact, what was used. However, rather than use the z-transform to design our filters, we obtained our coefficients directly from the b-plane triangle of stability. Figure 2.91 shows a three dimensional look at cutoff frequency as a function of (b_2,b_1).

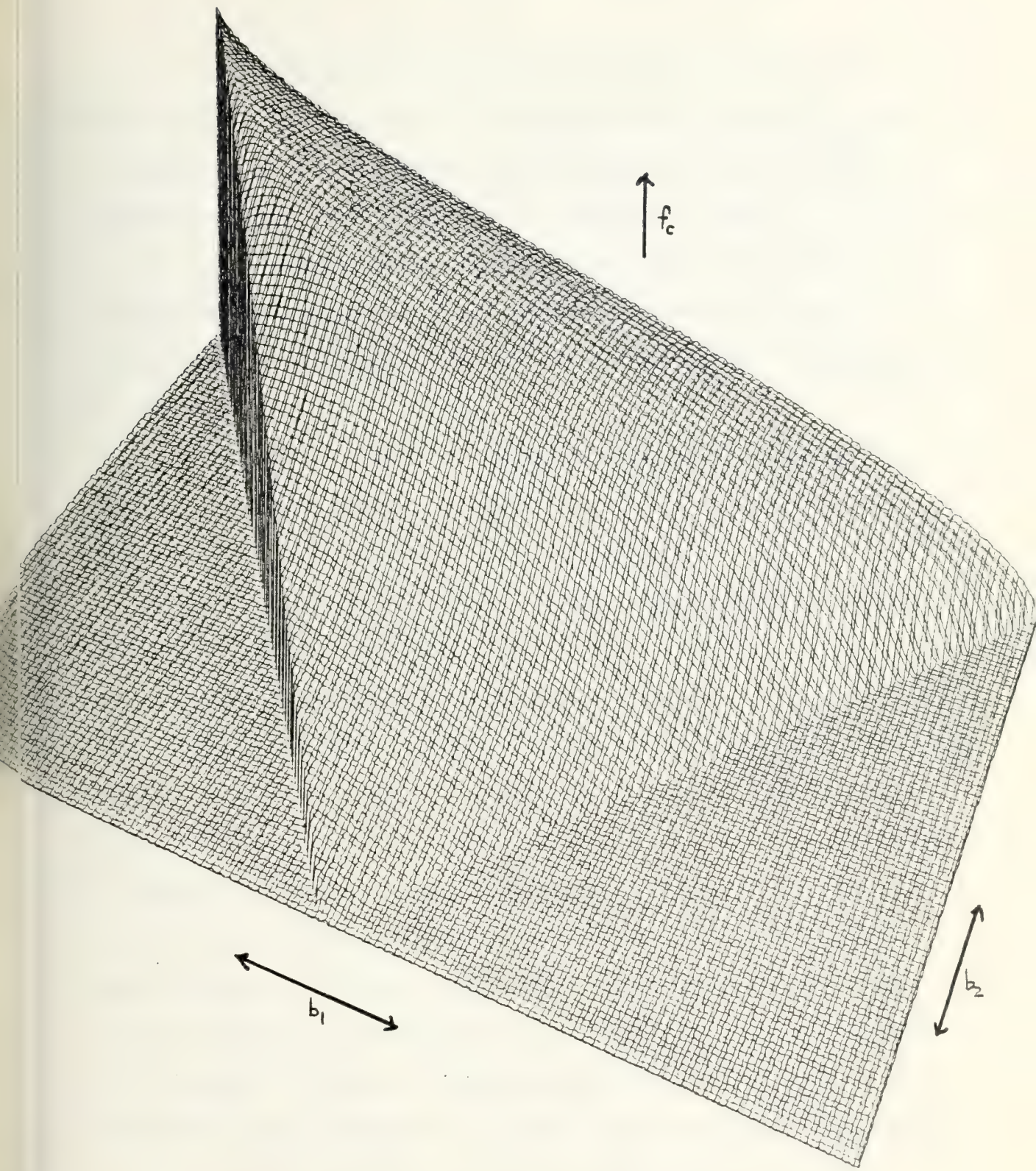
Section 2.6 Conclusions

Our goal when undertaking the experimental work was to show that sampled-analog devices, such as the CCD, were appropriate to the design of recursive comb filters. At the outset, we were one of the only groups working on this project, although many had utilized the charge transfer devices for delay elements in transversal filters. On the basis of the data contained in this chapter, we believe that we have shown that sampled analog devices can be used as the delay element in recursive comb filters.

There are several other conclusions which are significant:

A. The standard z transform is not very useful when designing to a frequency specification.

B. The bilinear z transform is much more useful, since it provides that the range of the sampled analog filter is precisely equal to the range of the analog filter from which it was derived.



ure?.91. Cutoff frequency (vertical scale) as a function of (b_2,b_1). The base
ne)-plane. The b_2 axis runs from foreground (b_2 = -1) to background (b_2 = 1).
b_1 axis runs across the page from +2 to -2 (left to right).

C. One can make the eight-to-first order and sixteenth-to-second order translations for a specific set of feedback and feedforward coefficients, thus affording the engineer with a powerful tool with which to design filters.

D. Inefficiency, the sample-and-hold effect, and the frequency response of the potentiometers (and the operational amplifiers, too) detract from perfect operation of the filters.

E. Second order sampled analog filter design is best done with a general purpose computer, using the program this research generated, or its equivalent.

F. Analog to Sampled analog second order filter design is not very tractable, and is not any better (in fact, worse) than using the computer program.

G. For the engineer at the bench, one easy way to predict the changes in the transfer function as a result of modifications of the coefficients is the b-plane triangle of stability curves. Cutoff frequency, ripple, the position of H_{max}, and the slope at cutoff have been calculated and isocontours are now available. Also, marginal changes in these quantities (especially cutoff frequency and slope) as a function of small incremental changes in b_1 and b_2 are now established and shown in four of the triangle drawings.

H. The eight bit CCD does not provide enough "teeth" at frequencies far below $f_s/2$ to make a true comb filter. The RETICON SAD device is better. Still better, at least insofar as number of delay stages is concerned,

is the Fairchild CCD (455 delay stages) recently acquired.

I. The magnitudes of the "null" frequencies in first and second order filters are about the same. Therefore, if this were the dominant specification, one would be advised to use first order filters for simplicity. On the other hand, second order filters make it possible to achieve higher selectivity, Q.

CHAPTER THREE

APPLICATION OF THE SAMPLED-ANALOG RECURSIVE COMB FILTER

3.0 Introduction

The contents of this chapter serve to supply the link between device/circuit research and systems applications. The primary application we have selected arises rather naturally in radar systems, that of the elimination of the echoes of stationary targets and the enhancement of the echoes of moving targets. Such a process is usually performed by a device called a Moving Target Indicator (MTI). The use of MTI is not novel ... navy air search radar systems have had designed-in MTI for twenty years. The problem was, however, that cancellation was generally done (not too well) by analog means and by use of a device called a quartz delay line. The physics of this device are adequately described in Skolnik's famous text, Introduction To Radar Systems. The following subjects will be addressed:

Section 3.1. Theory of echo cancellation; cancellers, filters

Section 3.2. Time domain truncation: effects on echo spectrum and cancellation

Section 3.3. The recursive comb filter as the MTI processing element

Section 3.4. Conclusions.

Section 3.1 Theory of Echo Cancellation; Cancellers, Filters

Echo cancellation and the distinguishing of stationary target echoes from moving target echoes is made theoretically possible by the physical phenomenon of doppler frequency shift. Briefly, doppler frequency shift occurs between a transmitted wave and a received echo wave when there is relative motion between transmitter and target. A discussion of doppler frequency shift, which can be observed not only in electromagnetic waves, but also in acoustic waves, can be found in most basic physics textbooks. If the relative motion between the two objects serves to decrease the distance between them, then the doppler frequency is considered positive, and the frequency of the received signal is higher than the transmitted frequency. Conversely, if the relative motion between the two bodies serves to increase the distance between them, then the doppler frequency is negative and the frequency of the received signal is lower than the transmitted frequency. If the distance between the objects is unchanged during the time of the transmission and reception of the pulse, then the frequency of the received signal is exactly the same as the transmitted frequency and the doppler frequency is said to be zero.

Since echoes from stationary and moving targets occupy different regions in the frequency domain, it is theoretically possible to distinguish between them ... to build equipments which will ignore stationary targets, but sound alarms on fast moving targets. After the echoes have been categorized

into stationary and moving targets, of course, the system can be made to produce almost any response ... including the automatic firing of a missile. But it is the initial problem of determining the radial speed of a target which has proven difficult over the years.

The doppler frequency is given by the following equation:

$$f_d = \frac{2\ v_r\ f_t}{c} \qquad (3\text{-}1)$$

where v_r is the radial speed of the target, f_t is the transmitted frequency, and, of course, c is the speed of light. It perhaps provides insight here to determine real doppler frequency for real parameters. The transmitted frequency of the AN/SPS-29 series air search radars is about 215 mHz. A target traveling at 400 knots of radial speed will produce a doppler frequency shift in this radar set of 296 Hz. Observe that the doppler frequency is about 0.000138% of the transmitted frequency. Since such a small percentage could rarely be detected by RF discriminators, the usual procedure is to shift the entire received echo spectrum to baseband, or so that the transmitted frequency is shifted to zero frequency. Then a bandpass filter could be used to distinguish the target echo at 296 Hz.

The analysis of doppler frequency has, of course, been done. Skolnik proceeds as follows:

A radar pulse is transmitted for $t_o < t < t_o + T$ and is given during that time by

$$V_{ref} = A \sin 2\pi f_t t \qquad (3\text{-}2)$$

The received signal is

$$V_{echo} = B \sin\left(2\pi(f_t+f_d) - \frac{4\pi f_t R}{c}\right) \qquad (3\text{-}3)$$

In the argument of V_{echo}, the first term represents the shift in frequency due to the doppler effect. The second term represents the shift, relative to V_{ref}, in phase of the echo due to the target's range.

Maximum unambiguous range is $R_{max} = cT/2$ where T is the time between pulses.

We extract the doppler information by mixing (multiplying) the V_{ref} by the V_{echo}. We want the difference frequency which will be of the form

$$V_o = C \sin (2\pi f_d t - \phi) \qquad (3\text{-}4)$$

Now there have been historically two major ways of processing video, or of presenting echo information to the operator. The first, chronologically, was the A-scope which provided to the operator a pulse at some distance along the horizontal axis which corresponded to the range of the target. MTI processing for the A-scope took the form of the well known "Butterfly" effect. Consider Equation (3-4). If the radial speed of the target is zero, then the doppler frequency is zero. Then we are left with

$$V_o = C \sin(-\phi) \qquad (3\text{-}5)$$

which is a constant for all time.

On the other hand, if the radial speed was not constant, then $f_d \neq 0$ and V_o oscillates in time. Thus, the operator, on observing an oscillating target echo on his A-scope knew with no further processing that the target was a moving target. On the other hand, he could safely ignore constant sized pulses since they represented stationary targets. Figures 3.1 and 3.2 show, respectively, the simulated A-scope presentation of a stationary target echo and a moving target echo.

The second form of display has been the popular PPI (plan position indicator) which is shown in Figure 3.3, and is well known to almost every naval officer. Here the plot is in terms of angle and range, not amplitude. Thus, the operator cannot distinguish between a constant-amplitude pulse and an oscillating-amplitude pulse, since the indication on the PPI is in terms of power. In order to provide some stationary vs. moving target processing, it is necessary to pre-filter the echoes prior to introduction to the PPI. The simplest method is the one zero canceller shown in Figure 3.4. Here the first echo from a target arrives at time t_o, and the second at time $t_o + T$. These signals are introduced into the subtractor with the following results:

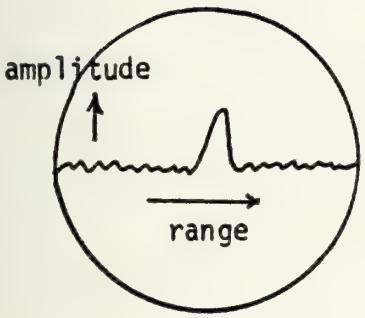


Figure 3.1. A-scope presentation of a stationary target. Notice the constant amplitude.

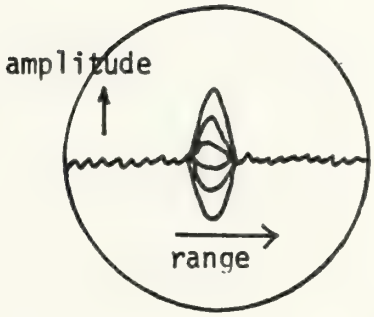


Figure 3.2. A-scope presentation of a moving target. Notice the so-called "Butterfly" effect.

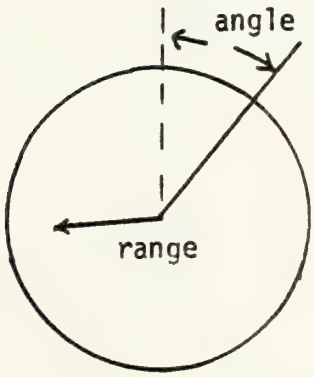


Figure 3.3. The plan position indicator (PPI) display.

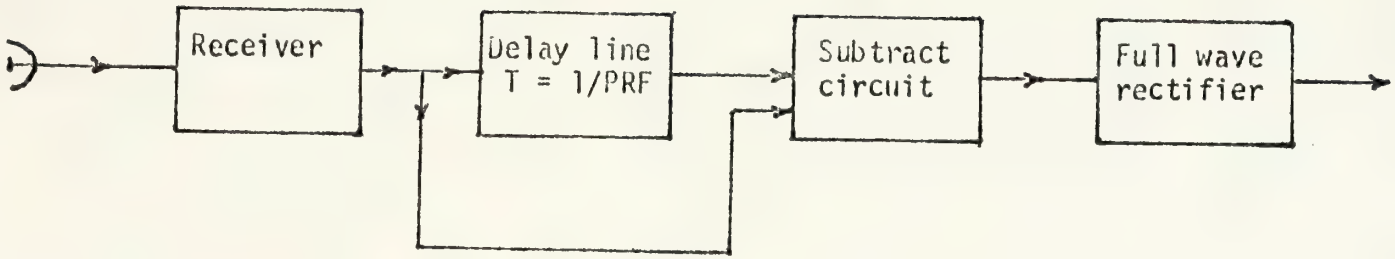


Figure 3.4. MTI receiver with delay line canceller (Skolnik)

$$V_{o1} = A \sin (2\pi f_d t - \phi)$$

$$V_{o2} = A \sin (2\pi f_d (t+T) - \phi)$$

$$V_{o1} - V_{o2} = k \sin \pi f_d T \cos (2\pi f_d (t+\tfrac{T}{2}) - \phi) \qquad (3\text{-}6)$$

As can be seen by an inspection of Equation (3-6), if the doppler frequency is zero, then the amplitude of the response for all time is zero. On the other hand, if the doppler frequency is not zero, then there is some nonzero amplitude response which is passed to the PPI.

It must, of course, be pointed out that there are many different doppler frequencies for which Equation (3-6) is zero for all time. Such doppler frequencies are given by

$$f_d = n/T \; ; \quad n = 0,1,2,... \qquad (3\text{-}7)$$

The meaning of this last equation is that if the target has a radial speed such that the doppler frequency satisfies Equation (3-7), then this processor will provide zero response. The physical interpretation is that the target will never be detected. Such speeds are called BLIND SPEEDS and many different designs have been tried to eliminate them, or at least reduce their effects. To determine how significant these blind speeds may be, an example with real parameters is in order. Referring to the aforementioned AN/SPS-29

series air search, we postulate the following characteristics:

$$f_t = 215 \text{ mHz}$$

$$PRF = 1/T = 266 \text{ Hz}$$

$$R_{max} = 250 \text{ mi.}$$

For these parameters, the first blind speed occurs at

$$v_n = c/2f_tT = 305 \text{ knots}$$

The first blind speed, then, is 305 knots, a speed well within the operating capability of many aircraft and missiles. As a matter of fact, the second and third blind speeds, 610 and 915 knots, also must be considered as possible target speeds. The significance, then, of blind speeds is great.

It must be emphasized that no amount of pulse shaping can eliminate the undesirable zero response at blind speeds. This phenomenon is caused by the periodicity of the target echoes, which, in turn, is caused by the periodicity of the radar transmitted pulse. One way of looking at the phenomenon which is appropriate to our discussion, is that the radar echoes represent a sampling of a signal waveform of the doppler frequency. Figure 3.5 is drawn to make this aspect of the phenomenon clearer.

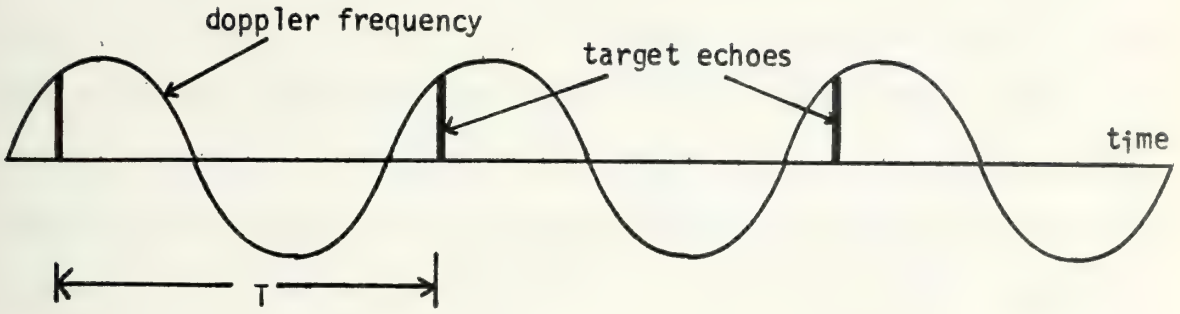


Figure 3.5. Radar encounters a target whose radial speed is such that the doppler frequency is exactly equal to the PRF. Since all target echo amplitudes are equal, the target provides the same input pulse train to the MTI as a stationary target. The response of the MTI filter is zero at this BLIND radial speed.

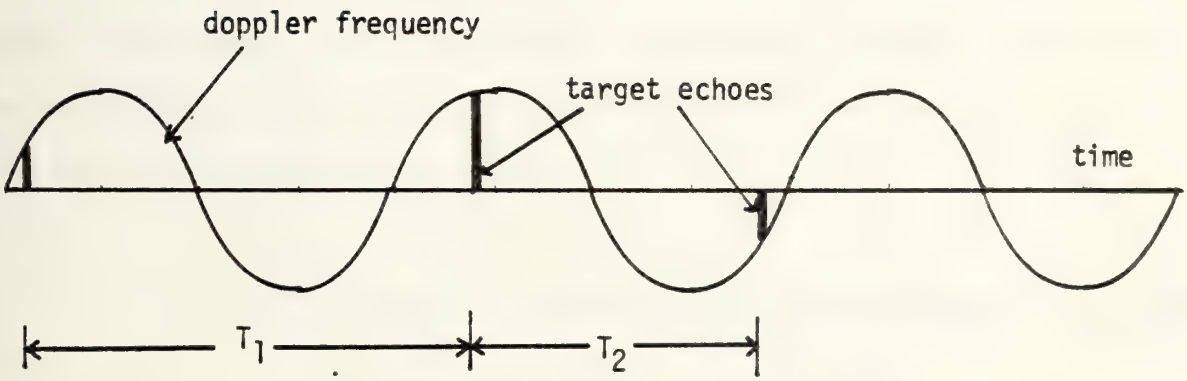


Figure 3.6. Radar encounters the same target (equal doppler frequencies) but with staggered PRF, the target echo amplitudes are not equal (as shown). Thus cancellation will not take place and the response of the MTI filter will be nonzero

One way of eliminating blind speeds is to change the pulse repetition frequency. Several methods of change have been proposed and used. One is called "staggered PRF" which utilizes two different PRF's and shifts them after each series of pulses, or sometimes, after each pulse. Complex gating in the radar set insures that the targets at equal ranges will be processed together in the MTI. The other method is called "Jittered PRF" which selects, at random, from a range of PRF's, a new PRF for each pulse. This method also requires complex gating to insure comparison of target echoes at appropriate ranges. The effect of staggered PRF is shown in Figure 3.6. The doppler frequency is drawn for all time and is sampled at those times during which it is illuminated by a radar pulse. Notice that the amplitudes of the target (traveling at a blind speed) echo are not equal, and thus, will not cancel completely in MTI canceller. As a matter of fact, the first blind speed now occurs at the first multiple such that

$$n/T_1 = m/T_2 \quad \text{where n and m are positive integers.}$$

T_1 and T_2 are the pulse repetition intervals respectively.

The traditional means of accomplishing analog delay has been the acoustic delay line. The steps involved in using this delay line are

1. Convert electromagnetic energy to acoustic energy.

2. Insert the acoustic energy into a quartz or mercury delay line ... which is a crystal shaped so that the energy will bounce from one face to another in a predetermined manner in order that it will appear at the output after a specified delay time.

3. Reconvert the acoustic energy output to electromagnetic energy.

The disadvantages of the analog delay have been known for sometime. They are

1. The delay line and associated transducers are quite lossy providing severe insertion loss.

2. The processor is bulky and large.

3. The crystal is sensitive to temperature variations. Changes in temperature change the shape of the crystal slightly causing a shift in the nominal delay time. Such shifts in delay time destroy proper operation of the canceller.

4. There are bandwidth limitations for some applications.

5. The crystal is sensitive to shipboard shock and vibration.

6. It is difficult to manufacture the crystal to required specifications ... or if it is done, it is extremely costly. Slight imperfections can cause secondary responses (ghosting).

7. Inhomogeneities in the material itself can cause spurious responses as well.

Today, many radar systems have digital processors. The raw analog target data is first converted to digital data for insertion into digital processors. Digital filters have been built and are utilized in MTI applications. Even some recursive filters are utilized in this way. Of course, in any analog to digital system, there are certain disadvantages:

1. Analog to digital conversion produces some inevitable quantization noise ... so-called noise resulting from the rounding off of the actual analog values. This quantization noise takes on considerable significance when using recursive filters since, in recursive filters, the same digital data is fed back to the input repeatedly thus increasing the quantization error. A brief example will show what we mean. Suppose we have a feedback coefficient of +0.9 and we insert a unit pulse into the filter: Round off the output to the nearest tenth.

Start with input at time zero = 1.0 ... output = 1.0
time one = 0.9 ... output = 0.9
time two = 0.81 ... output = 0.8
time three= 0.72 ... output = 0.7
time four = 0.63 ... output = 0.6
time five = 0.54 ... output = 0.5
time six = 0.45 ... output = 0.5

but no matter how much further we go, we cannot change the output from 0.5. In this extremely simple example, we have

shown that the quantization error can be quite significant in digital systems. The steady state error in our system was the difference between zero and half the original value of the signal, clearly unacceptable.

Of course, there are known ways of combatting such catastrophic errors. The first is to expand the number of quantization levels. As more quantization levels are provided, the error is consequently reduced. Of course, this means more hardware, and greater complexity of design. Perhaps of greater import than the actual hardware is the computer study which must be undertaken to determine exactly how many more quantization levels are required. Several theses and papers originating at the Naval Postgraduate School have led the field in this subject (see especially, work by Dr. S. R. Parker). These computer studies must be made by radar/processing engineers and, thus, are costly.

Another way of reducing the effect of digital quantization error is to prevent the samples from being multiplied over and over by the same quantized gain coefficient. This takes the form of eliminating the canonical for digital filters and substituting transversal filters, or, at least, using cascaded recursive sections of not more than two orders each. Quantization error in digital systems can be reduced, but it cannot be eliminated, and we are just beginning to understand methods to predict significant quantization error and take countermeasures.

2. Making the conversion from digital back to analog is also a source of noise.

3. Naturally, the hardware involved in each of these conversions increases the complexity of the equipment, and also increases the design time of the radar system itself.

There is one other important factor to notice about these cancellers. Since they are filters, they have a transient response to any input signal. As with all linear filters, their frequency transfer functions are the Fourier transforms of their unit pulse responses. A more useful specification for these cancellers is the so-called indicial function which is nothing more than the response of a linear system to a unit step input. The relation between the unit pulse response and the indicial function is

$$h(nT) = A(nT) - A((n-1)T) \qquad (3\text{-}8)$$

where h is the unit pulse response and A is the indicial function. Ideally, then, we would want the indicial function, A, to be identically zero since the echo from the stationary target is tantamount to receiving a unit step input. Unfortunately, the indicial function is not zero for all time, but has a nonzero transient. This transient, of course, can be calculated for any given filter, but it is often erroneously ignored, and it must be considered when assessing the value of a filter. We have measured some of the time

responses in the laboratory of various cancellers. The measurements agree quite closely with the predicted values of h and A. Figure 3.7 shows two cases of the use of filters which have zeros in their transfer function at dc, and yet the indicial function and unit pulse response are nonzero for a substantial time period. For the case in Figure 3.7a, the feedback coefficient of our first order filter was only -0.3 and the indicial function dies out rather rapidly. The magnitude of the function falls below 0.1A(0) at n = 5 and stays below that value. On the other hand, the filter with the feedback of -0.9 has an extended response such that A does not fall below 0.1A(0) until the 29th pulse output.

Naturally recursive filters are not unique in having nonzero transient behavior. Transversal filters (cancellers) also exhibit this troublesome characteristic. In the canceller shown in Figure 3.4, the transient response to a unit step input can easily be calculated and is shown in Figure 3.8. The transient response in this case is minimal, but, of course, it is accomplished at the cost of a nonideal frequency characteristic which is shown in Figure 3.9. The passbands of the filter are clearly not flat, a circumstance which can, at times, lead to a loss of detection. This can arise when the operator notices that the level of the video is too high for a target traveling at a radial speed for which the response of the filter is maximum. The operator's impulse is to reduce the gain of the radar set, and when he

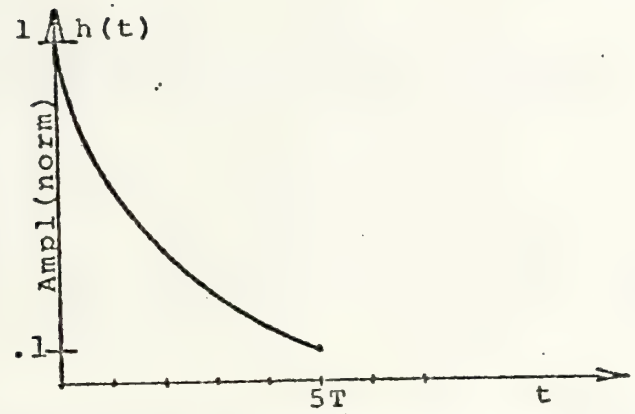


Figure 3.7a. Unit pulse response of first order recursive filter, $b_1 = 0.3$

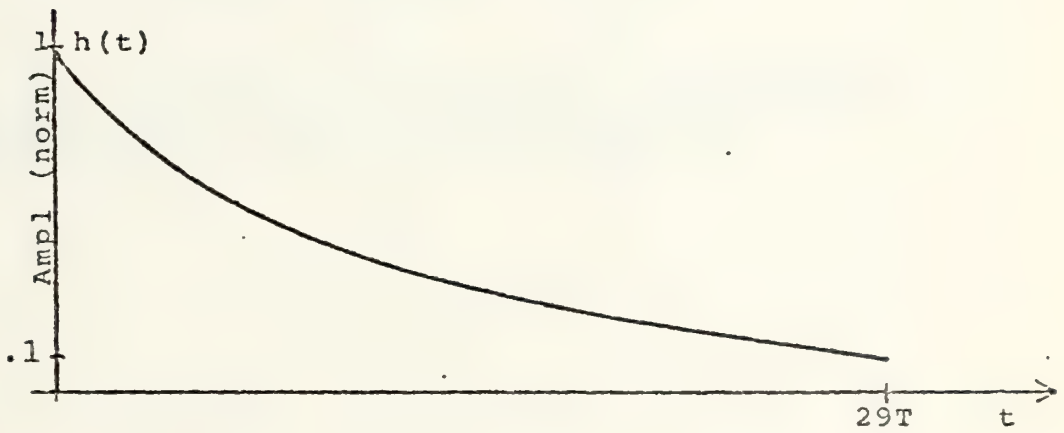


Figure 3.7b. Unit pulse response of first order recursive filter, $b_1 = 0.9$

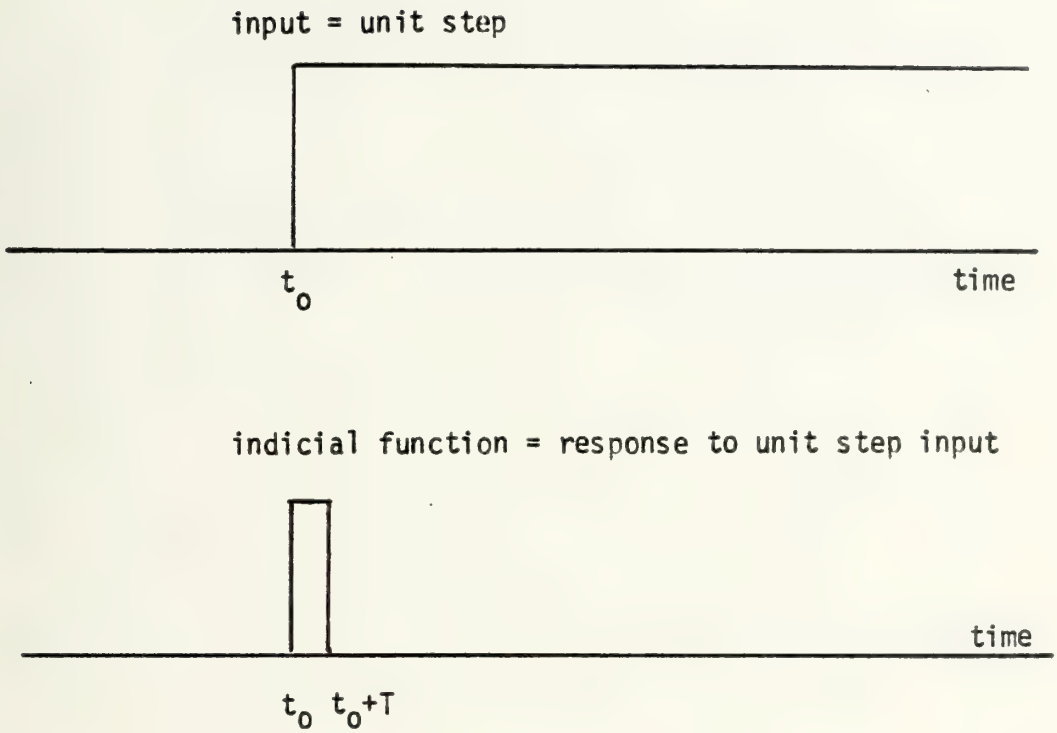


Figure 3.8. Transient response of one zero-no pole MTI canceller to unit step input. Filter will exhibit another pulse when input falls to zero. Filter is shown in figure 3.4.

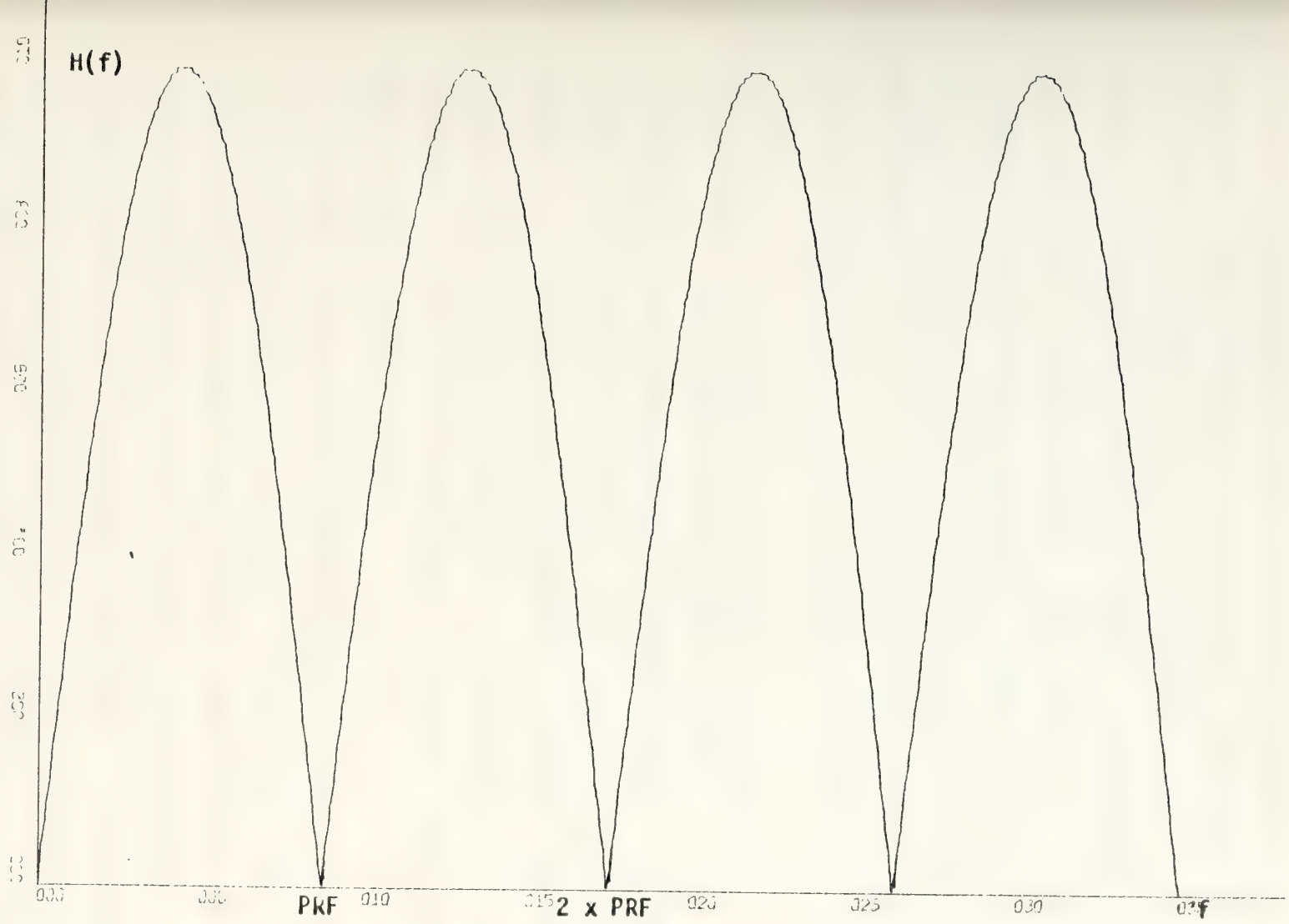


Figure 3.9. Transfer function of first order (one zero, no pole) canceller for MTI.

does so, he eliminates the possibility of detecting targets at some of the marginal radial speeds. Flatness in the pass-band can, of course, be better achieved by increasing the order of the filter, but this implies lengthening the unit step response. As can be seen from Chapter two, and particularly the second order case, filters close to the ideal frequency characteristic can be designed. Filters whose time characteristic is close to the ideal can also be designed. But filters "close" to both, cannot.

Section 3.2 Time Domain Truncation: Effects on Echo Spectrum and Cancellation

There is one deviation from theory which is so significant in radar systems that it deserves a section all by itself. It is not often considered in designing a radar system with MTI, but its effects are far reaching and important. It is the time domain truncation that almost every search radar exhibits.

Consider a normal air search radar. It is mounted high on a ship and rotates in azimuth. Pulses are transmitted and are returned from targets before the rotation of the antenna can move far enough to avoid receiving the echoes. If this were not so, no echoes would ever be received. The significant development is that the antenna only stays "on target" for a small finite time. It will be shown that this small finite time is extremely important in determining the response of the MTI and, thus, the ability of the radar to distinguish stationary from moving targets.

The small finite time of "on target" enables the radar to send out k pulses, each of similar width. It is this number, k, which turns out to have such a large effect on the spectrum of the target echo train and the subsequent cancellation.

As k approaches infinity, of course, we have the ideal case which we had assumed in Section 3.1. The spectrum of the stationary target is periodic (spikes at dc, and each multiple of the PRF). The spectrum of moving targets is also periodic with period equal to the PRF. Clearly, the filter with the frequency characteristic having zeros at dc and the multiples of the PRF will eliminate the stationary target completely and pass the moving target echoes. Figures 3.10a and 3.10b show the ideal frequency spectrum for stationary and moving targets.

As k is reduced from infinity, we no longer have the ideal case. Physically, this represents the antenna scan increasing or the pulse repetition frequency decreasing. The spectrum of the input, then, can no longer be discrete, although it will still be periodic. This follows from the familiar properties of the Fourier transform pairs: If a function is not periodic in one domain, then it will be continuous in the other domain.

Since the spectrum has become continuous (and if it is not everywhere zero), then there must be a band of frequencies in which substantial signal energy lies. If we choose some

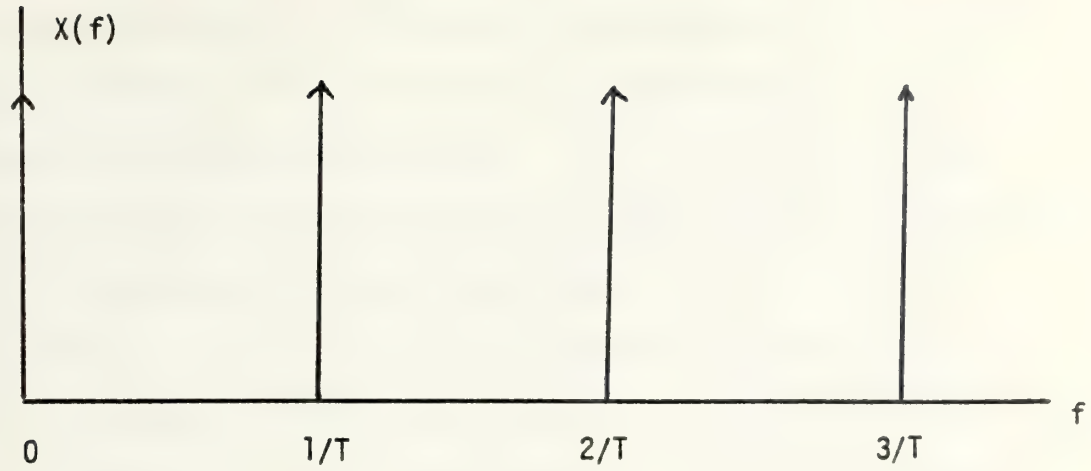


Figure 3.10a. Ideal spectrum of the target echoes from a stationary target. Spacing of the lines in the spectrum is 1/T (PRF).

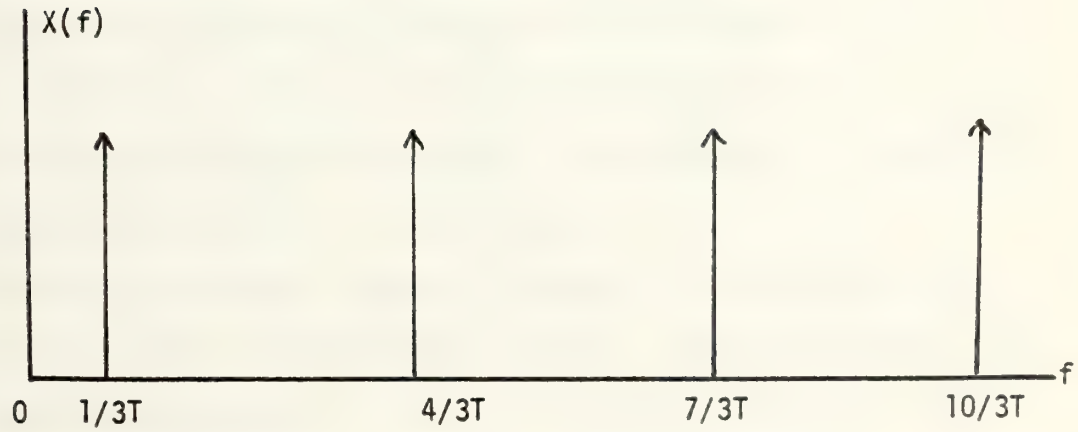


Figure 3.10b. Ideal spectrum of the target echoes from a moving target. Target radial speed is such that the doppler frequency is $f_d = 1/3T$. The spacing of the lines is 1/T (PRF).

k, we can calculate the frequency characteristic directly. We have selected three values of k to demonstrate the effect of time domain truncation: For k = 2, Figure 3.11 shows that the energy in the signal (still from a stationary target) is hopelessly spread throughout the band from dc to 1/T. For k = 6, we can see in Figure 3.12 that the energy is confined to the major lobes at dc and 1/T, but still a substantial amount of energy exists in the nominal passband of some MTI filter. For k = 10, Figure 3.13 shows that the level of the minor lobes has fallen from the preceding case and that most of the energy lies in the region around dc and 1/T. The relative amount of energy of the signal in the nominal passband of the MTI filter compared to that in the stopband can, of course, be calculated for any given filter and any given number of "hits on target".

Of course, the moving targets are affected by time domain truncation just as the stationary targets are. Their spectra are also spread throughout the passband and stopband with the attendant loss of efficiency.

We have calculated several examples to help provide an insight into these truncation effects. The assumptions for our calculations are:

1. Speed of moving target is such that the response of the MTI processor is at a maximum.

2. PRF is fixed at $f_r = 1/T$.

3. Pulse width is identical for moving and stationary targets and is given by τ.

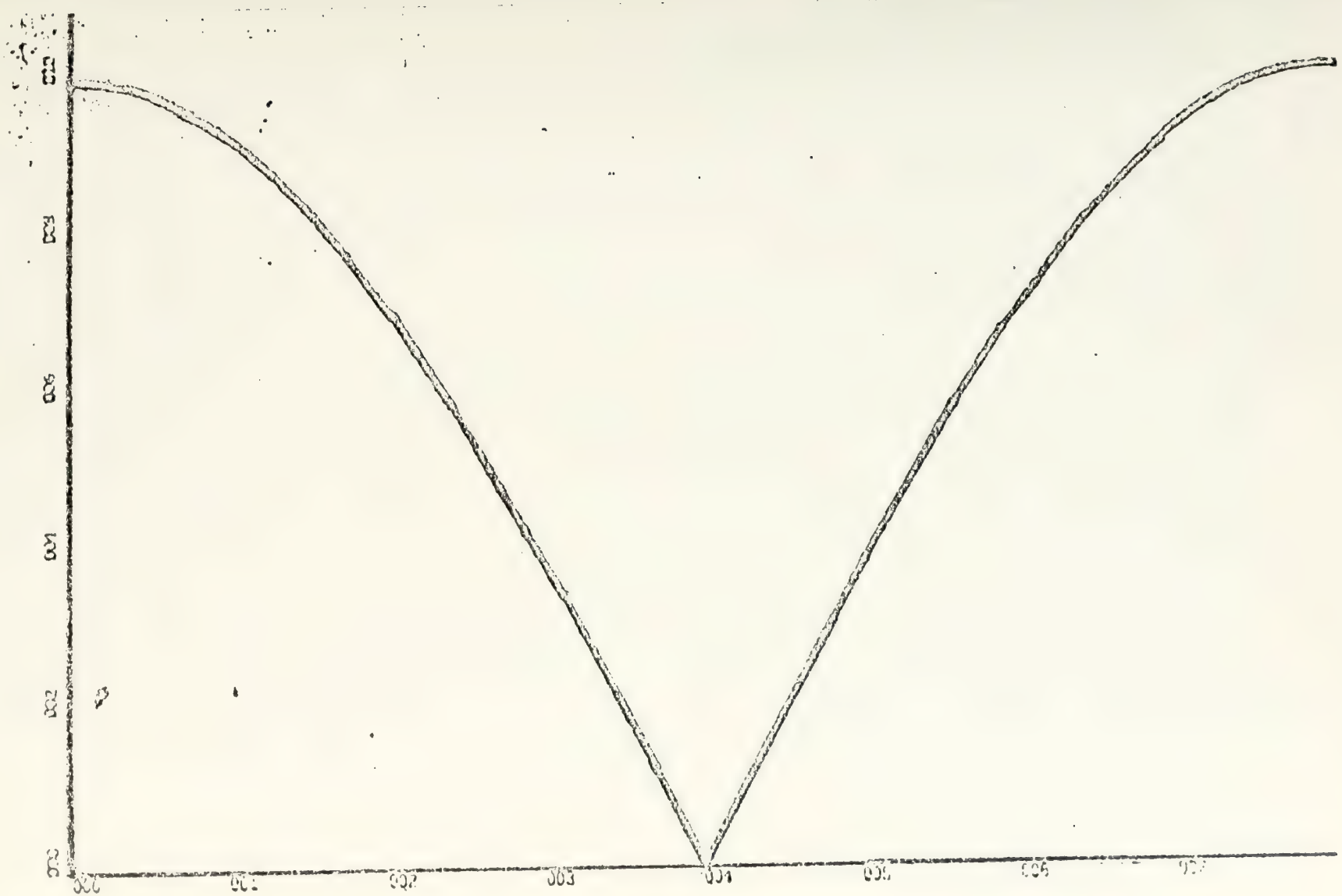


Figure 3.11. Frequency spectrum of stationary target where echo pulse train consists of k = 2 hits per scan

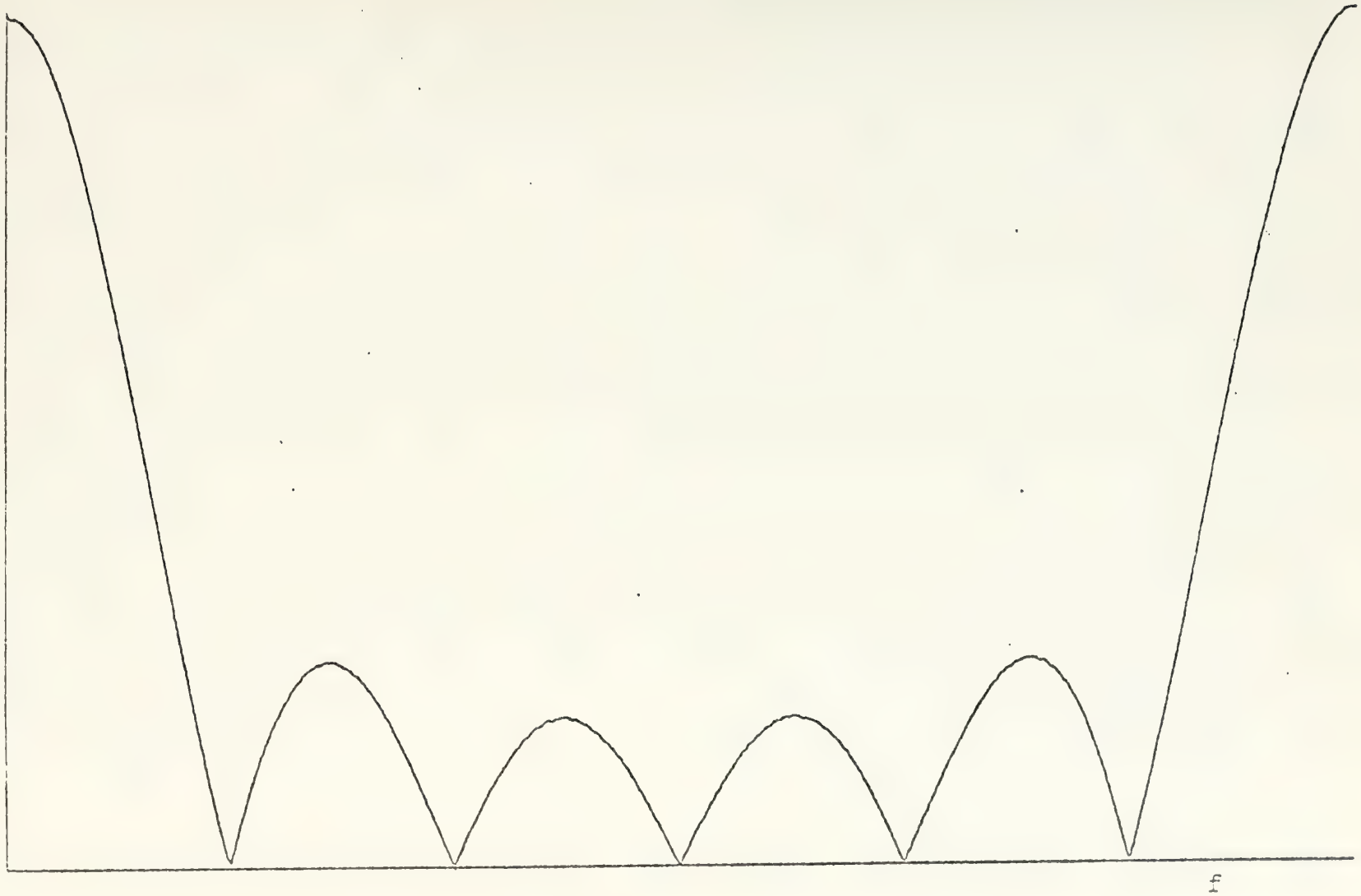


Figure 3.12. Frequency spectrum of stationary target where echo pulse train consists of k = 6 nits per scan

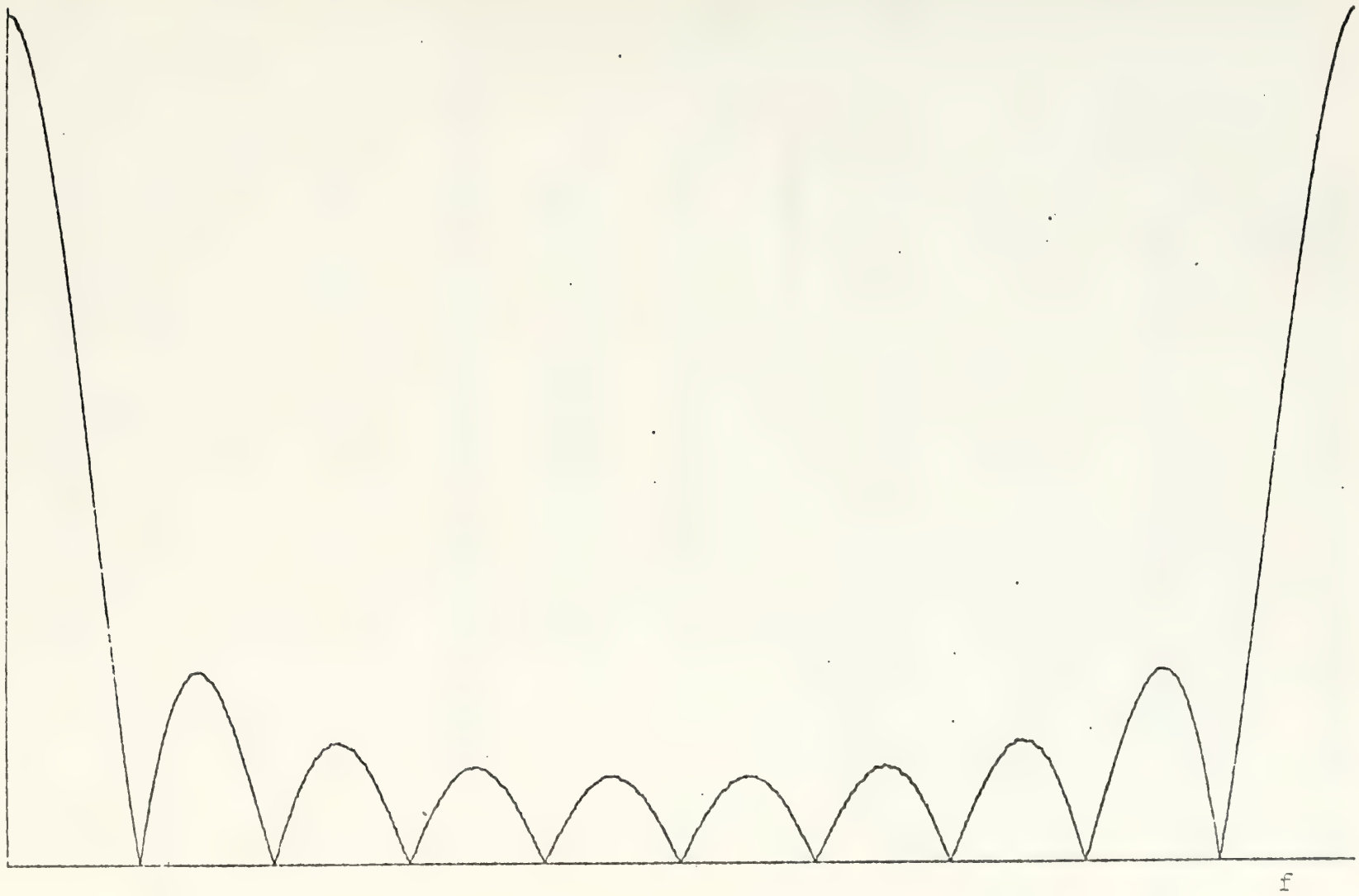


Figure 3.13. Frequency spectrum of stationary target where echo pulse train consists of k = 10 hits per scan

4. There is equal input energy from stationary and moving targets.

First we calculate the output spectrum based on the nonideal input spectrum and the frequency transfer function of the filter:

$$Y_1(f) = H(f)\ X_1(f) \tag{3-9}$$

where Y_1 is the output spectrum for an input X_1. X_1 is defined as the nonideal input spectrum of a truncated stationary target echo train. $H(f)$ is the transfer function of the filter.

Also we have, for moving targets:

$$Y_2(f) = H(f)\ X_2(f) \tag{3-10}$$

Following this calculation, we form the integral

$$E_1 = \int_0^{1/T} |Y_1(f)|^2\ df \tag{3-11}$$

and

$$E_2 = \int_0^{1/T} |Y_2(f)|^2\ df \tag{3-12}$$

Finally, we form a ratio of the above quantities. Let

$$R = \frac{E_2}{E_1}\ ; \qquad R_{db} = 10 \log \frac{E_2}{E_1} \tag{3-13}$$

Thus R is the measure of energy contained in the output from an input, truncated, moving target echo train relative to the energy contained in the output caused by a truncated, stationary target echo train, with the four assumptions made above.

In the ideal case, this ratio would be infinite. In the nonideal case (nonideal, yet still tractable since the behavior can be predicted), the ratio will never be infinite. The ratio will form what is an upper bound for the radar system under these conditions. Under no circumstance, can the user expect to see a greater difference between output moving target echo energy and output stationary target echo energy than R (again, under the aforementioned assumptions).

The results of these calculations are tabulated and shown in Table 3.1.

As might be expected from an intuitive approach, the larger feedback enables the filter to achieve a higher signal-to-noise ratio (where noise is defined as the stationary target echo). Also, as expected, the ratio rises with the number of hits on target. There are no real surprises here, but it is emphasized, that prior to the design of a radar MTI processor, one must decide on many specifications. One of the important specifications is the expected number of "hits" on target. The problem is tractable either for a "least number of hits" expected, or some mean number

TABLE 3-1

ENERGY RATIO: Energy contained in the output of an MTI processor caused by a moving target (truncated pulse train) compared to the energy contained in the output caused by a stationary target (truncated pulse train). This figure represents the maximum "Signal-to-Noise" ratio possible given the assumptions listed in the text. Ratio is in decibels.

	First Order Feedback Coefficient		
NUMBER OF HITS	0.3	0.7	0.9
2	6.7	12.3	15.9
10	15.3	19.3	23.5
20	18.5	22.9	27.0

of hits. But, most importantly, the problem must be faced at the beginning of design, not, as in some cases related to us, after construction of the radar when the best answer is a retrofit which never really works.

As an aside, it has occurred to us that since we do know the "noise" spectrum and signal spectrum, that it is possible to design an optimum filter (maximizes the signal-to-noise ratio). Such an optimum filter can be designed using the derivation given by Taub and Schilling [31] for nongaussian noise:

$$H_{opt}(f) = X^*(f)/G(f) \qquad (3\text{-}14)$$

where $X^*(f)$ is the complex conjugate of the input signal spectrum, $G(f)$ is the noise energy spectrum, and $H_{opt}(f)$ is the transfer function for the optimum filter. Figure 3.14 shows the computer-aided optimum transfer function.

This transfer function is admittedly very limited. It is only optimum for the four assumptions given above. If the target changes radial speed, then the filter is no longer optimum, a fact which probably precludes its use. On the other hand, optimum filters could be designed for a whole bank of doppler frequencies (radial speeds) and the output of each compared after each input echo. Such a scheme, while theoretically sound, is probably more than is really needed in the field of MTI.

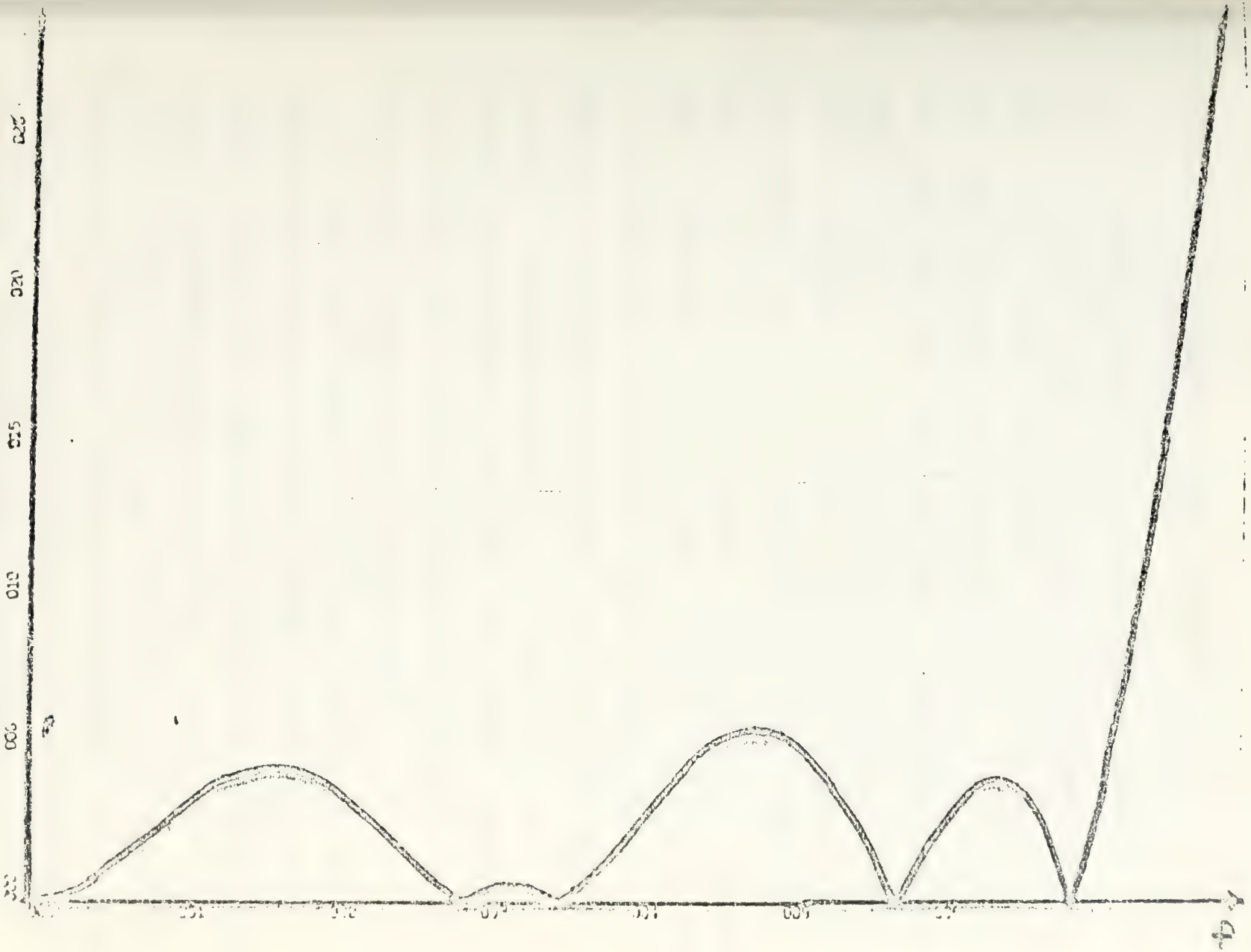


Figure 3.14. Transfer function of optimum filter for moving target whose doppler frequency is half the PRF vs. stationary target.

Section 3.3 The Recursive Comb Filter as the MTI Processing Element

The original AN/SPS-58 air search radar utilized a digital recursive filter as its main MTI processor. The OPEVAL of this radar was conducted in 1972. In 1975, the radar set was still undergoing redesign. One of the problems associated with this radar was the difficulty that was experienced with second-time-around radar targets. A second-time-around target is a target (moving or stationary) whose echo appears in the pulse repetition interval succeeding the one in which it should appear. That is, the echo appears after transmission of the succeeding pulse. The echo appears as if it were at a small range from the radar, when in actuality, it is at a great range. The actual ambiguity is equal to the maximum range of the radar. If, for instance, the echo appears such that the apparent range is, say 50 miles, and the maximum range of the radar is 250 miles, then the actual range is 300 miles.

One means to eliminate, or at least categorize, second-time-around echoes is to change the PRF and see if the echo changes its position drastically. Changing the PRF naturally changes the maximum range of the radar. If the echo is a first-time-around echo — a legitimate echo — then changing the PRF will not affect its position on the PPI scope. However, if it is a second-time-around echo, then changing the PRF such that the maximum range of the radar is, say 270 miles, will shift the position of the echo by 20 miles so

that its apparent range is 30 miles. When the target "jumps" like this, the operator is trained to conclude that his target is really a second-time-around echo.

The OPEVAL of the AN/SPS-58 (Westinghouse, prime contractor) continued to have severe problems with second-time-around echoes into 1975. The reason given for NOT providing the operator with a PRF dither was

> "Strong land clutter does break through at the top of the display at a range of about 19 miles but the MTI performance is very good. The recursive digital MTI is not amenable to quick PRF changes so that PRF dither and range correlation for the removal of second-time-around targets is not a feature of this radar."*

The reason for believing that one could not change the PRF with a recursive digital filter is a long-standing tradition among radar designers, who believe that the length of the transient is sufficient to interfere with cancellation, and that it is not desirable to provide PRF changes in a recursive filter. Part of this feeling stems from a reading of Skolnik's excellent text. In it, he warns that second-time-around echoes cannot be cancelled by PRF changes ... that, as a matter of fact, staggered PRF's should be avoided if second-time-around echoes from stationary targets are a problem. His analysis is thorough and

---

*Private communication with Mr. James Whitaker, NELC, Code 1260 who quoted official documents dealing with the AN/SPS-58 OPEVAL of 18 July 1972.

correct. However, the interpretation on his words is sometimes inaccurate. For instance, there is NO reason at all to avoid changing the PRF with a recursive filter so long as the timing circuits in the MTI processor are adjusted so that one always compares echoes at the same range. This synchronization must be performed anyway (no matter what form of filter is used) whenever PRF "dither" or staggering is done. This concept can become clearer by the following example. Figure 3.15 shows a transmitted pulse train where, for simplicity, between each pair of transmitted pulses, there are ten range gates. These range gates occupy a certain range band (if the maximum range is 100 miles, then each range gate contains the energy received from a range band of 10 miles).

After the second transmitted pulse, the maximum range is changed to 80 miles, but ten range gates still are present. Now, by simple division, we find that the range band width is only 8 miles.

Since there are ten range gates, we shall be using an imaginary filter with ten delay elements (a 10th order filter). After the first pulse is transmitted, echoes (or lack thereof) for the ten range gates are received and stored successively in the delay elements. After the second pulse is transmitted, the contents of these range gates are consecutively compared with the new information (the echoes from the second pulse) and we get, serially, ten

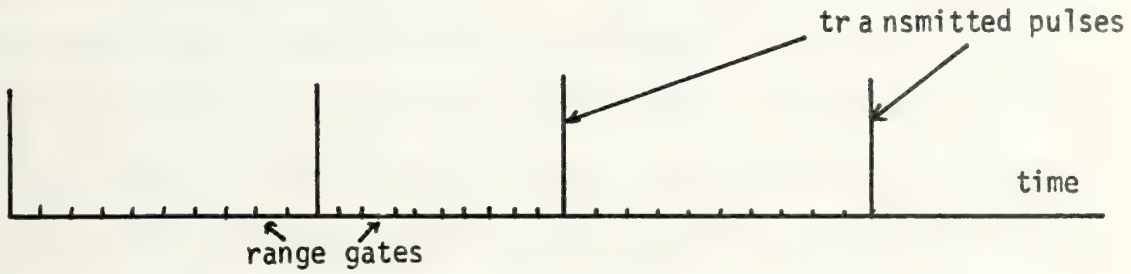


Figure 3.15. Transmitted pulse train with staggered PRF. Ten range gates of information introduced to MTI after each pulse.

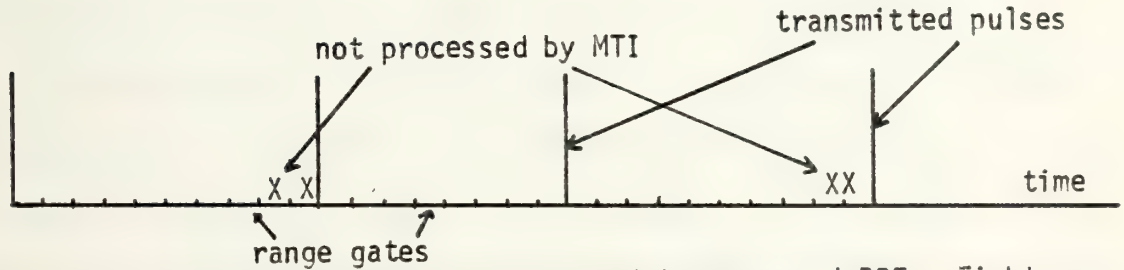


Figure 3.16. Transmitted pulse train with staggered PRF. Eight range gates (of equal range-width) of information introduced to MTI after each pulse. The farthest ranges are not processed by MTI.

outputs ... one from each range gate. In the sixth range gate, for instance, we are comparing data from the range band of 50 miles to 60 miles. The output is a function of the two echoes received from that band of ranges.

Now the pulse repetition interval is changed for the third pulse. There are still ten delay elements and ten range gates, but each range band is only eight miles wide. The output from the filter as the echoes from the third pulse arrive no longer makes any sense. The reason is that when we compare the contents of the sixth range gate with the arriving echo, we are comparing energy in the 50-60 mile range to energy in the 40-48 mile range for the third pulse. No conclusions can be drawn from such a comparison.

But this type of misalignment can exist in transversal filters too. Depending on the order of the filter, the effect is either serious or tolerable. It is well to be reminded that the longer the transversal filter is, the more closely it resembles a recursive filter.

This alignment is rather easily avoided. Most of the time, the MTI processing is not required on range bands far from the transmitter. In this case, it is a relatively easy engineering task to synchronize the filter to begin each processing phase after each pulse, and to maintain range gate width no matter what PRF is used. Figure 3.16 shows the conceptual pulse train for this plan. Notice that there are ranges for which no processing in the MTI takes place. This turns out to be acceptable in most applications.

The frustration in the design of the AN/SPS-58, at least as far as MTI is concerned need not to have been so great. Yet, the idea of recursion in a filter was such an anathema, that PRF changes weren't even attempted. Mr. J. W. Taylor, leader of the Westinghouse effort on the project, spoke on the subject of specifications:

> "Invariably, intense concentration on any single problem produces an optimum design which actually increases the vulnerability to some other problem. The ideal radar receiver is one which offers adequate protection from all sources of interference rather than maximum rejection of any one." **

Mr. Taylor was deeply concerned with the problems associated with designing an MTI processor which was "too good" since, in his view, it prevented the radar from dealing with other problems of detection, specifically, the second-time-around problem. We believe that a PRF dither WAS feasible in the radar design and that it could have easily dealt with the second-time-around problem.

One other approach to using recursive filters was done by Westinghouse in a subsequent design. The recursive filter utilized was a fourth order filter with interrupted operation. That is, the filter was operated normally for twenty-seven pulses, then blanked for five pulses. This had the effect

---

** Private communication with Mr. James Whitaker, NELC, Code 1260 who quoted official documents dealing with the AN/SPS-58 OPEVAL of 18 July 1972.

of dumping (or almost dumping) the contents of each delay element in the filter. Thus when the next section of twenty-seven pulses begin, the filter was empty. Such a filter has been termed a "truncated filter". Several truncated filters of this type were investigated and transfer functions were derived. The truncated filters, as a class, could be characterized by having time variant feedforward and feedback coefficients. When the coefficients go to zero, the output goes to zero. At first, this seemed like a good way to proceed, but subsequent results indicated that, since the coefficients were time-varying, the transfer function was also time-varying, and the reaction of the filter to a stationary or moving target was a function of the time the echo was presented to the filter. Transient behavior was increased, since the filter was artificially restarted from time to time. The approach did not turn out to be profitable.

Simulation Results:

Using the charge transfer device manufactured by RETICON corporation, the serial analog delay (SAD), we devised a method to simulate the action of the filter on a train of echo pulses. Figure 3.17 shows the block diagram of the simulation. The square wave generator was used to sample the input pulse trains. The sinewave generator simulated a moving target while the function generator provided dc and simulated a stationary target. It was possible to

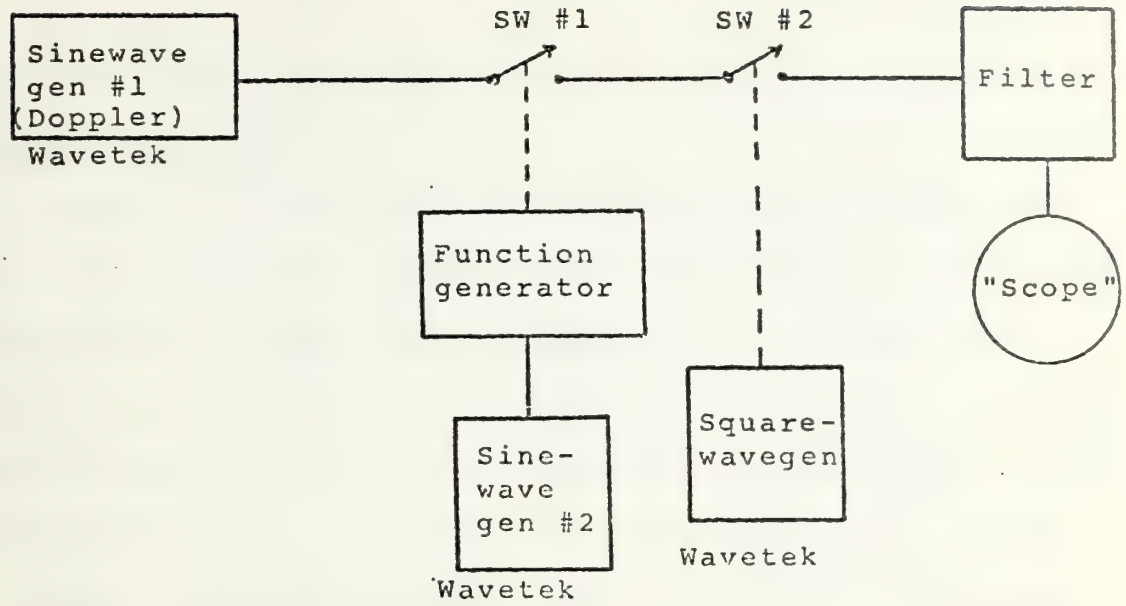


Figure 3.17. Block diagram of simulation configuration.

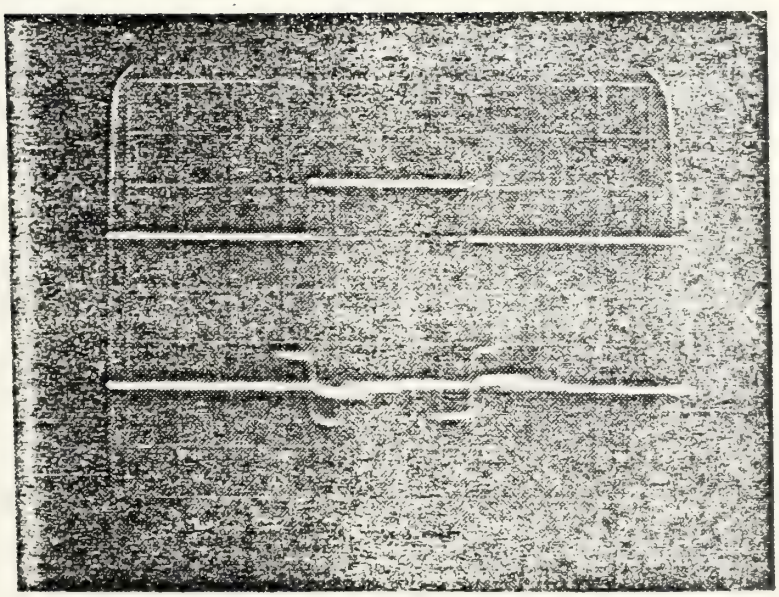


Figure 3.18. Cancellation of "stationary target", Horizontal: 10 μsec/cm. Vert: 0.5 V/cm

inject into the filter both a stationary target echo and a moving target echo at the same time. Such an approach was used with good results.

Figure 3.18 shows the cancelling of a stationary target by our recursive filter. The upper trace of the oscilloscope shows the input echo. The lower trace shows the transients associated with the beginning and the end of the input pulse, but the region between the ends is almost zero. We believe that we have achieved acceptable cancellation. In Figure 3.19, we see cancellation of a narrow stationary target. The transients do not have time to die out here so that the output never falls to zero.

Figure 3.20 shows the effect of the filter on a moving target. This oscillograph shows an A-scope presentation, so that we get the "Butterfly" effect on the output, as well as the input. One can see that the amplitude of the output is about the same as the input indicating that the insertion loss is quite small.

Figure 3.21 shows the effect of the filter on a combined moving target echo and stationary target echo. Although the quality of the picture is poor, cancellation of the stationary target echo was done quite well, while the moving target echo was faithfully passed. Figure 3.22 shows the same set-up with a different PRF (PRF = 10 kHz). Here can be seen quite clearly the cancellation of the dc component and the passage of the ac component (represented by the "Butterfly" effect).

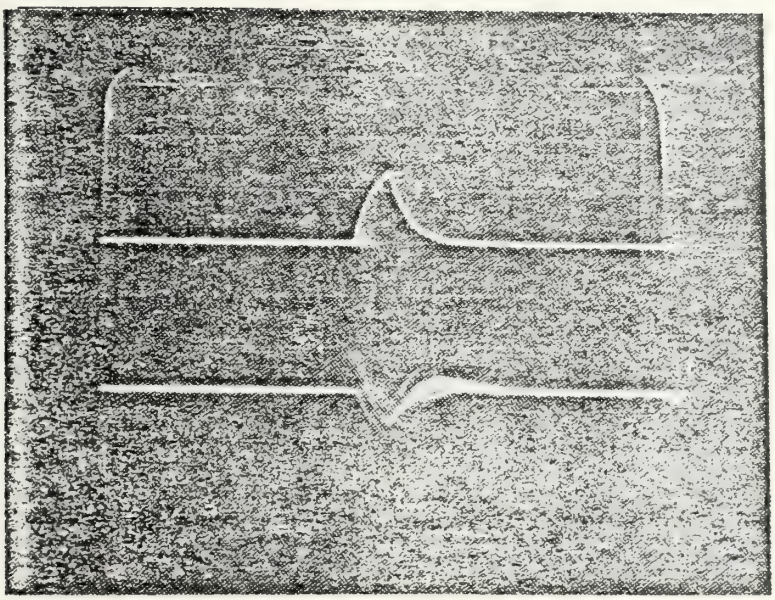


Figure 3.19. Cancellation of narrow, stationary target. PRF = 800 Hz. Horizontal: 10 μsec/cm.

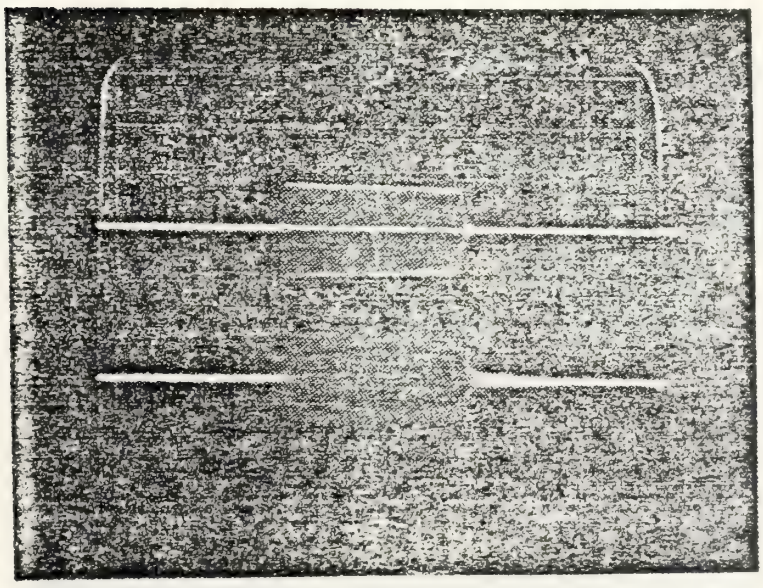


Figure 3.20. Moving target signal. PRF = 800 Hz. Horizontal: 10 μsec/cm.

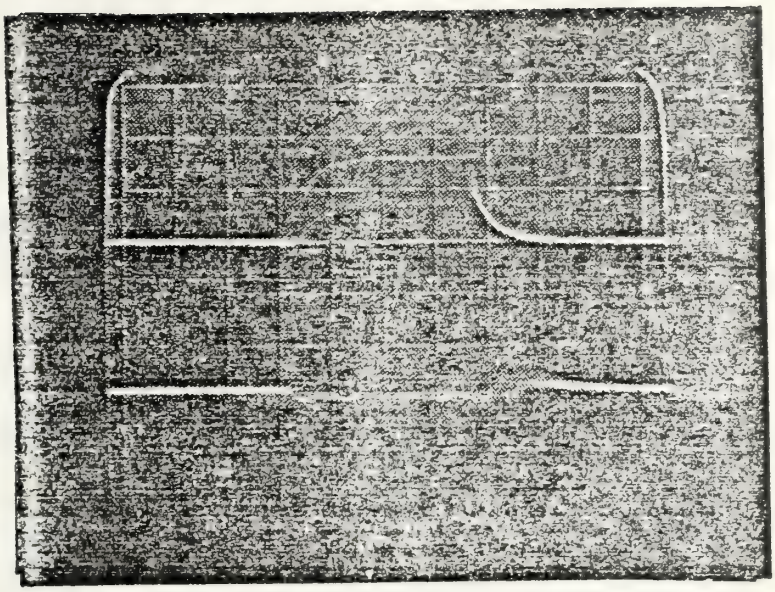


Figure 3.21. Moving and stationary targets. PRF = 800 Hz
Horizontal: 10 μsec/cm.

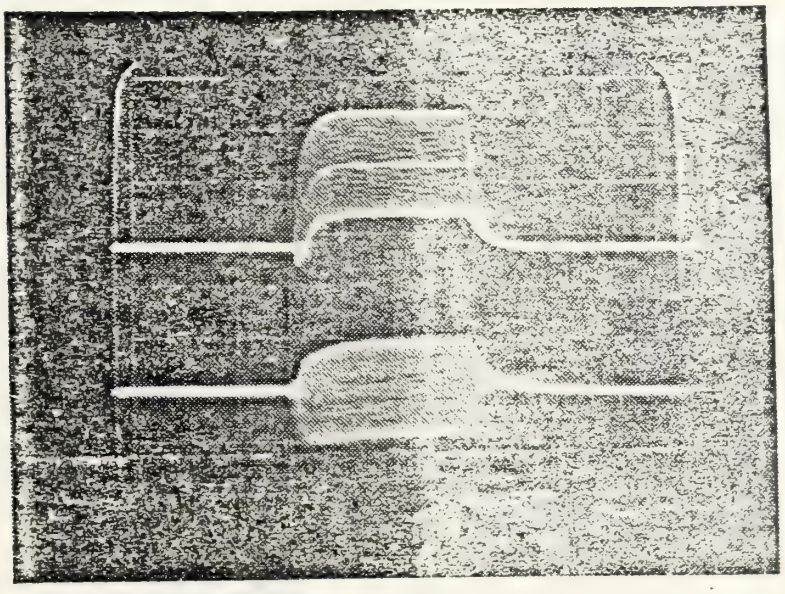


Figure 3.22. Moving target and stationary target,
f_d = 5 Khz, PRF = 10 Khz. Horizontal: 50 μsec/cm.

Figure 3.23 shows the interesting case when the doppler frequency is the same as the PRF. Theoretically we should get 100% cancellation and our simulation shows that, indeed, we do get good cancellation, although, qualitatively, not so good as with dc.

Section 3.4 Conclusions

We believe that we have established that recursive, sampled analog comb filters can be used in radar systems as viable MTI processors. Certainly, the analog nature of the signal relieves us of the A-D/D-A conversions and of the need to project quantization error. Moreover, for any given clutter spectrum, we can derive the appropriate transfer function. Following this determination, we can consult Chapter two to find the appropriate filter which offers the best fit to our ideal transfer function. Or, alternatively, we can apply the computer program contained in Appendix E and discussed in Section 2.4 and find the optimum filter relative to some cost function.

Our CCD which we used for much of the analysis in Chapters one and two was not suitable for radar systems work. The bandwidth required for radar work is approximately given by the half inverse of the pulse width:

$$BW = \frac{1}{2\tau} \qquad (3\text{-}15)$$

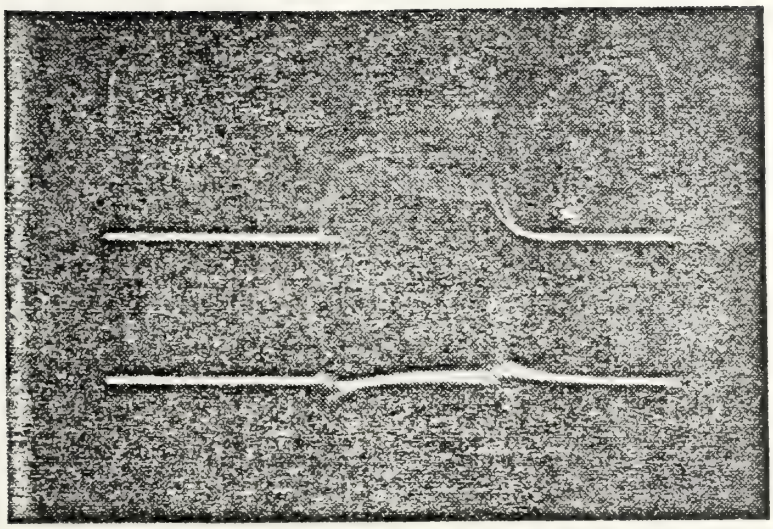


Figure 3.23. Moving and stationary targets superimposed. f_d = 10 Khz, PRF = 10 Khz. Horizontal: 50 μsec/cm.

The CCD had eight delay stages. The necessary clock frequency was

$$f_c = PRF \times 8 \tag{3-16}$$

A typical value of PRF is 800 Hz. Thus the clock frequency is 6400 Hz. This clock frequency is quite small and the CCD did not operate well at such low frequencies due to the noise produced by the dark current.

If the clock frequency is 6400 Hz, then the largest frequency for which no aliasing takes place (in accordance with Nyquist's sampling theorem) is 3200 Hz. Applying Equation (3-15), we find the pulse width:

$$\tau = \frac{1}{2(3200)} = 156 \ \mu sec \tag{3-17}$$

Such a wide pulse width is totally unrealistic. Since the clock frequency is tied to the PRF, we must have longer delay registers in order to increase the bandwidth without disturbing the clock frequency.

For the above reason, actual measurements with radar were not taken. Nevertheless, we feel that subsequent work with the 455 delay stage devices which the solid state laboratory has recently acquired will prove to be fruitful.

Time domain truncation is an important phenomenon and must be considered when designing the filter for an MTI.

Ideally, of course, it should be one of the major specifications determined prior to design. In practice, probably design and firming of the specification could take place concurrently. But it must not be ignored.

Section 4.0 Conclusions

Detailed conclusions appear at the end of each chapter. Overall conclusions are really no different from each set contained there. It appears, that with the development of the CCD's, we shall be able, in a short time to advance the state of the art to such an extent that we will be able to use recursive (and transversal) sampled analog comb filters to perform functions which have been performed by digital or analog means in the past. The sampled analog device has the advantage of no quantization error, no D-A/A-D hardware, easy execution of major functions (addition, integration, delay, etc.), and, at least potentially, low cost and low power consumption. The present sampled analog device has the disadvantage of too many bias voltages, drifting parameters, and the requirement for external level shifters to make it compatible with most of the other circuit elements.

On the other hand, the 455 stage CCD has the projected capability in a filter for achieving 227.5 "teeth" between dc and $f_s/2$ ... Neither the sample-and-hold circuit, nor the inefficiency will detract very much from the shape of the first few "teeth". Of course, the sampling frequency must be made sufficiently large so that the "teeth" of the comb are of the required bandwidth.

APPENDIX A

TRANSLATION FROM EIGHTH ORDER TO FIRST ORDER

Consider a transfer function of the form

$$H(z) = \frac{a_o + a_8 z^{-8}}{1 + b_8 z^{-8}} \qquad \text{(A-1)}$$

We have asserted in Section 2.1 that such a transfer function can be regarded as a first order function with suitable adjustments in sampling frequency, A graphical approach was used in Section 2.1. Here is an analytical approach:

Consider the poles only: The poles of the eighth order system lie on the z-plane at the positions determined by solving the following equation:

$$z^8 + b_8 = 0 \qquad \text{(A-2)}$$

or

$$z^8 = - b_8$$

Now

$$z^4 = \pm(-b_8)^{1/2} = b_8^{1/2} \angle 90°,\ b_8^{1/2} \angle 270°$$

and

$$z^2 = b_8^{1/4}\angle 45°,\ b_8^{1/4}\angle 225°,\ b_8^{1/4}\angle 135°,\ b_8^{1/4}\angle 315°$$

and

$$z = b_8^{1/8}\angle 22.5°,\ b_8^{1/8}\angle 202.5°,\ b_8^{1/8}\angle 112.5°,\ b_8^{1/8}\angle 292.5°,$$
$$b_8^{1/8}\angle 67.5°,\ b_8^{1/8}\angle 247.5°,\ b_8^{1/8}\angle 157.5°,\ b_8^{1/8}\angle 337.5°$$

For example, suppose b_8 is positive. Then the "eighth" order pole ($z^8 = -b_8$) lies on the negative real axis of the z-plane. The square root of this quantity yields two roots, each lying on the imaginary axis of the z-plane. The square root of these poles yields four roots each lying on the $b_8^{1/4}$ circle with angles $\pi/4$, $3\pi/4$, $5\pi/4$, $7\pi/4$. Finally, the square root of these four quantities yields the final result. The eight poles appear on the $b_8^{1/8}$ circle and on radial lines at $\pi/8$, $3\pi/8$, $5\pi/8$, ..., $15\pi/8$. It remains to show that the transfer function, then, is periodic with respect to $f_s/8$ as well as periodic with respect to f_s. This can be done by forming the equation

$$1 + b_8 e^{j2\pi f 8T} \stackrel{?}{=} 1 + b_8 e^{j2\pi\left(f + \frac{1}{8T}\right)8T} \qquad \text{(A-3)}$$

This equation reduces to

$$e^{j2\pi f8T} \stackrel{?}{=} e^{j2\pi\left(f + \frac{1}{8T}\right)8T}$$

or

$$e^{j2\pi f8T} \stackrel{?}{=} e^{j2\pi f8T}\, e^{j2\pi} \qquad \text{(A-4)}$$

Now since the second term in the argument of the expression on the right corresponds to a complete revolution in phase around the unit circle of the z-plane (that is, the $j2\pi$ term does not change the value of the expression), we have

$$e^{j2\pi f8T} = e^{j2\pi f8T} \qquad \text{(A-5)}$$

which is an identity.

Extension of this procedure to values of f beyond $f + f_s/8$ is easily done by observing that the only effect of multiplying $f_s/8$ by an integer is to produce some multiple of 2π in the argument of the expression on the right. As such, the value remains the same.

This proves that the transfer function given in Equation (A-1) is periodic with respect to $f_s/8$ for all real frequencies.

APPENDIX B

DETAILED CALCULATIONS FOR CRITICAL VALUES OF b_1 FOR FIRST ORDER FILTERS

The transfer function of a sampled analog filter designed by using the standard z transform is given in Section 2.2.1 (Equation (2-14)):

$$H(z) = \frac{1}{1 - e^{-2\pi f_x T} z^{-1}} \qquad \text{(B-1)}$$

The requirement for the integrator is that the cutoff frequency be less than $0.25f_s (0.25/T)$. The calculation of the cutoff frequency proceeds as follows:

A. Determine the gain of the filter at dc:

$$H(o) = \frac{1}{1 - e^{2\pi f_x T}} = \frac{1}{1 - b_1}$$

B. Square the gain and divide by two:

$$|H(o)|^2 = \frac{1}{2(1 - 2b_1 + b_1^2)}$$

C. Set the resulting quantity equal to the squared magnitude of the transfer function at the cutoff frequency:

$$|H(f_c)|^2 = \frac{1}{2(1 - 2b_1 + b_1^2} = \left| \frac{1}{1 - b_1 e^{-j2\pi f_c T}} \right|^2 \qquad \text{(B-2)}$$

The expression on the right must be squared using complex mathematics:

$$\left| \frac{1}{1 - b_1 e^{-j2\pi f_c T}} \right|^2 = \frac{1}{1 - 2b_1 \cos 2\pi f_c T + b_1^2} \qquad \text{(B-3)}$$

which gives us

$$\frac{1 + b_1^2 - 4b_1}{-2b_1} = \cos 2\pi f_c T \qquad \text{(B-4)}$$

Now we wish f_c to be less than $0.25f_s$. The upper limit for f_c occurs when $f_c = 0.25f_s$. Using this value, we solve for b_1:

$$1 + b_1^2 - 4b_1 = 0 \qquad \text{(B-5)}$$

or

$$b_1 = 2 \pm \sqrt{3} = 0.268 \qquad \text{(B-6)}$$

The other root, 3.723, clearly is unacceptable since b_1 must be less than unity for stability purposes.

Checking the exponential tables, if $b_1 = 0.268$, then the exponent $\omega_x T = 1.32$. This leads us to

$$f_x/f_s = 0.21 \tag{B-7}$$

which is the value used in Section 2.2.1.

The calculation for the critical value of b_1 such that if b_1 assumes any larger value, no cutoff frequency exists, follows. Start with Equation (B-4):

$$\frac{b_1^2 - 4b_1 + 1}{-2b_1} = \cos 2\pi f_c T \tag{B-8}$$

The limit of the cutoff frequency, because of the symmetry of digital and sampled analog filter transfer functions, is $f_s/2$. To insert this value into the right side of Equation (B-8) makes the value of the argument equal to π. Thus we have

$$b_1^2 - 4b_1 + 1 = 2b_1 \tag{B-9}$$

Solving this quadratic, we obtain two roots, but one is greater than unity and must be discarded because of the stability criterion. The other value is

$$b_1 = 0.172$$

Thus

$$-e^{-2\pi f_x T} = -0.172$$

and

$$2\pi f_x T = 1.76$$

and

$$f_x/f_s = 0.28$$

Thus, the critical value for b_1 is 0.172, and this corresponds to a value of 0.28 for the ratio between the analog filter cutoff frequency and the sampling frequency. Thus, if the analog filter cutoff frequency is 10 kHz, the sampling frequency must be at least 10,000/0.28 = 35.7 kHz in order for the sampled analog filter to have a cutoff frequency at all.

APPENDIX C

TRANSLATION FROM SIXTEENTH ORDER TO SECOND ORDER

Consider a transfer function of the form

$$H(z) = \frac{a_o + a_8 z^{-8} + a_{16} z^{-16}}{1 + b_8 z^{-8} + b_{16} z^{-16}} \qquad \text{(C-1)}$$

We have asserted in Section 2.3 that such a transfer function can be regarded as a second order function with suitable adjustments in sampling frequency.

Consider the poles only: The poles of the sixteenth order system lie on the z-plane at the positions determined by solving the following equation:

$$1 + b_8 z^{-8} + b_{16} z^{-16} = 0 \qquad \text{(C-2)}$$

or

$$z^{16} + b_8 z^8 + b_{16} = 0 \qquad \text{(C-3)}$$

The quadratic form can first be used to find expressions for z^8. Thereafter, successive square roots can be extracted to obtain, ultimately, the expressions for the positions of the poles:

The quadratic form, applied to the above equation, yields

$$z^8 = \frac{-b_8}{2} \pm \frac{b_8^2 - 4b_{16}}{2} \qquad \text{(C-4)}$$

or, for simplicity

$$z^8 = a \pm jb = R\angle\theta\ ,\ R\angle{-\theta} \qquad \text{(C-5)}$$

Carrying on, we find

$$z^4 = R^{1/2}\angle\theta/2,\ R^{1/2}\angle\theta/2+\pi,\ R^{1/2}\angle{-\theta/2},\ R^{1/2}\angle{-\theta/2+\pi}$$

and

$$z^2 = R^{1/4}\angle\theta/4,\ R^{1/4}\angle\theta/4+\pi,\ R^{1/4}\angle\theta/4+\pi/2,$$
$$R^{1/4}\angle\theta/4+3\pi/2,\ R^{1/4}\angle{-\theta/4+\pi},\ R^{1/4}\angle{-\theta/4+\pi/2},$$
$$R^{1/4}\angle{-\theta/4},\ R^{1/4}\angle{-\theta/4+3\pi/2}$$

Finally, the sixteen poles are located at

$$z = R^{1/8}\angle\theta/8,\ R^{1/8}\angle\theta/8+\pi/2,\ R^{1/8}\angle\theta/8+\pi,\ R^{1/8}\angle\theta/8+3\pi/2$$
$$R^{1/8}\angle\theta/8+\pi/4,\ R^{1/8}\angle\theta/8+5\pi/4,\ R^{1/8}\angle\theta/8+3\pi/4,$$
$$R^{1/8}\angle\theta/8+7\pi/4,\ R^{1/8}\angle{-\theta/8},\ R^{1/8}\angle{-\theta/8+\pi},$$
$$R^{1/8}\angle{-\theta/8+\pi/2},\ R^{1/8}\angle{-\theta/8+3\pi/2},\ R^{1/8}\angle{-\theta/4+\pi/4},$$
$$R^{1/8}\angle{-\theta/4+7\pi/4},\ R^{1/8}\angle{-\theta/4+3\pi/2},\ R^{1/8}\angle{-\theta/4+7\pi/4}$$

Figures C.1, C.2, C.3, and C.4 show successive square root extractions. Figure C.1 shows the position of the two values for z^8. Figure C.2 shows the position of the four values of z^4, C.3, the eight values of z^2, and, finally, C.4, the sixteen values of z.

From Figure C.4, we observe that the poles are not equidistant from each other as they were in the eight-to-one order translation. Nevertheless, they do lie at positions which are symmetrical with respect to dc, $f_s/8$, $2f_s/8$, ..., $7f_s/8$. This fact is sufficient to show that the transfer function is periodic with respect to $f_s/8$, but it rests on a graphical method. The analytical approach is as follows:

Prove that

$$1 + b_8 e^{-j2\pi f8T} + b_{16} e^{-j2\pi 16fT} = 1 + b_8 e^{-j2\pi (f + \frac{+}{8T}) 8T} + b_{16} e^{-j2\pi (f + \frac{1}{8T}) 16T} \tag{C-6}$$

This equation reduces to

$$b_8 e^{-j2\pi f8T} + b_{16} e^{-j2\pi f16T} = b_8 e^{-j2\pi f8T} e^{-j2\pi} + b_{16} e^{-j2\pi f16T} e^{-j4\pi} \tag{C-7}$$

As in the case for translation from eighth order to first order, use is made of the fact that a phase term in the argument of a multiple of 2π does not affect the value of

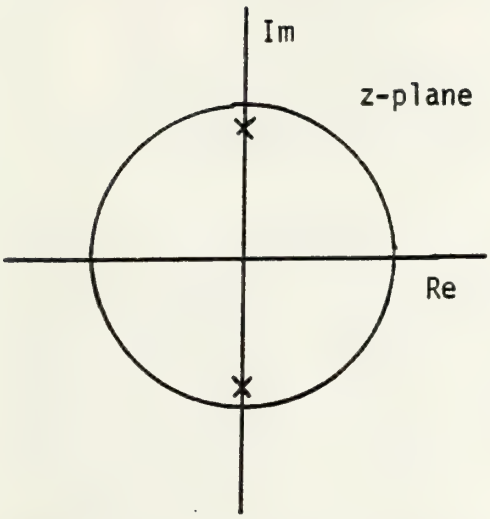


Figure C.1. The two values for z^8. Each happen to be purely imaginary.

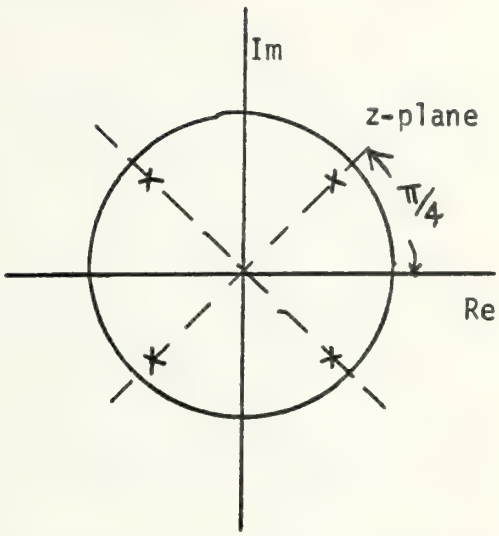


Figure C.2. The four values for z^4.

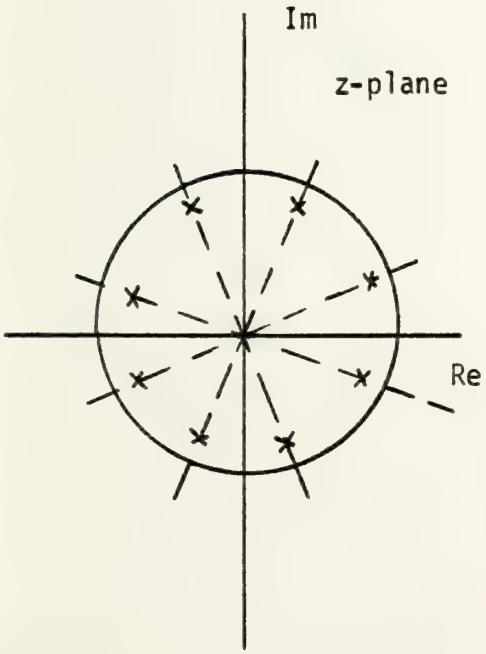


Figure C.3. The eight values for z^2.

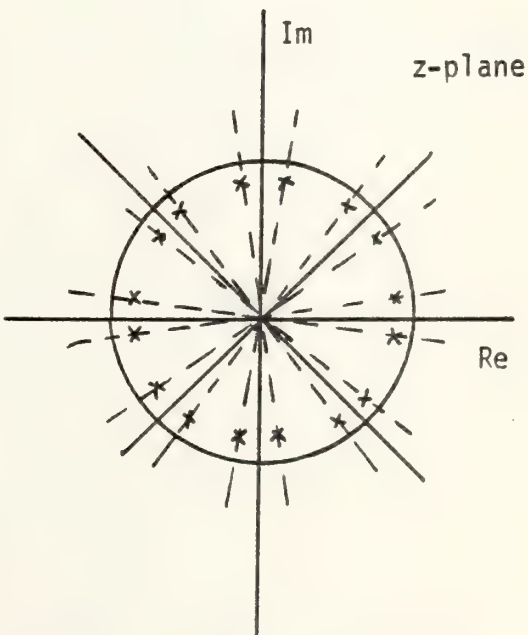


Figure C.4. The sixteen values for z. Note that they are symmetric about the radial lines $nf_s/8$ where n=0,1,2,...,7.

the exponential expression. Thus we have

$$b_8 e^{-j2\pi f8T} + b_{16} e^{-j2\pi f16T} = b_8 e^{-j2\pi f8T} + b_{16} e^{-j2\pi f16T} \quad \text{(C-8)}$$

which is an identity.

Extension to values of frequency greater than $f_s/8$ is easily made by observing that the only change occurs as an integral multiple of 2π in the argument of the expression on the right.

This proves that the transfer function given in Equation (C-1) is periodic for real frequencies with respect to $f_s/8$.

APPENDIX D

COMPUTER PROGRAM FOR DETERMINING THE CUTOFF FREQUENCY OF A SECOND ORDER SAMPLED ANALOG FILTER

C This program requires no data as such. The feedback coefficients, b_1 and b_2, take on their entire range of permissible values. The program is written for

$$a_o = 1$$
$$a_1 = 2$$
$$a_2 = 1$$

which is a low pass filter transfer function. The a's may be changed by changing the coefficients in the numerator for the expression for H(K) in statement 19. Several other variables are dimensionalized, but not used. These variables are used in other programs where the cutoff frequency is also needed.

```
MAIN PROGRAM
      DIMENSION H(251),FC(99),SL(99),SB1(99),SB2(99)
C          Normalize the sampling frequency
      T = 1.
      PI=3.141592654
C              The next statement begins the major DO loop of the program.
      DO 1 I=1,199
      XI=I-1
      B1=-1.98+XI*0.02
C              B1 takes on values from -1.98 to 1.98 at intervals of 0.02
               The next DO loop defines the value of b2
      DO 2 J=1,99
C              Initialize the significant variables
      FC(J)=0.
      SB2(J)=0.
      SB1(J)=0.
      SL(J)=0.
      XJ=J-1
      B2=-0.98+XJ*0.02
C              b2 takes on values from -0.98 to 0.98 at intervals of 0.02
C              The following statements test to see whether the b1,b2 pair
C              is outside the triangle of stability or not.  If the pair
C              is outside, then FC(J) remains equal to zero and the next
C              value of b2 is tried.
      IF(B1-B2-0.999)4,2,2
    4 IF(B1+B2+0.999)2,2,5
C              If the pair is within the triangle of stability, DO loop
C              number three is entered:
    5 DO 3 K=1,251
      XK=K-1
```

```
C               Frequency is defined.
      W=2.*PI*0.002*XK*T
C               The transfer function is defined for W
      H(K)=(6.+8.*COS(W)+2.*COS(2.*W))/(1.+B1**2+B2**2+2.*B1*COS(W)+2.*
     1B2*COS(2.*W)+2.*B1*B2*COS(W))
      IF(H(K))14,15,15
C               H(K) should not be negative since in its computation we have
C               multiplied a complex number by its complex conjugate.  Nevertheless,
C               in single precision work, H(K) can become slightly negative.
C               When this happens, the following statement of square root
C               extraction obviously cannot be done.  Thus, if the calculated
C               H(K) is negative, it is arbitrarily set to zero.
   14 H(K) = 0.
   15 H(K)=SQRT(H(K))
      IF(XK)6,6,7
C               Find the maximum value of H
    6 HMAX=H(K)
      GO TO 3
    7 IF(H(K)-HMAX)3,3,9
    9 HMAX-H(K)
    3 CONTINUE
C               Determine the ripple (defined in accordance with the
C               material in Section 2.4)
      RIP=(HMAX-H(1))/2.
      C=0.5*((H(1)+RIP)**2)
C               Find the H(K) which is closest to the square root of C.
C               It is this K which defines the cutoff frequency
      K=251
   13 CONTINUE
      IF((H(K)**2)-C)11,11,12
   11 K=K-1
      GO TO 13
C               Interpolation
   12 R=((H(K)**2)-C)/((H(K)**2)-(H(K+1)**2))
      XK=K-1
C               Determination of the cutoff frequency
      FC(J)=0.002*XK+R*0.002
    2 CONTINUE
      PRINT 49, B1
      PRINT 50,(FC(J),J-1,99)
    1 CONTINUE
   49 FORMAT(1X,'B1=',F8.4)
   50 FORMAT(1X,13E10.4)
      STOP
      END
```

APPENDIX E

A COMPUTER PROGRAM FOR CALCULATING THE OPTIMUM COEFFICIENTS FOR A SAMPLED ANALOG FILTER USING CASCADED SECOND ORDER SECTIONS ... programmed in Fortran

```
MAIN PROGRAM
      IMPLICIT REAL*8 (A-H,O-Z)
C          All real variables will be double precision
      EXTERNAL FUNCT
      DIMENSION H(184),X(16),G(16)
      COMMON/RAW/W(100),Y(100),M,ICALL
C     REad in Data
      WRITE (6,51)
   51 FORMAT ('1INPUT DATA')
C          The following is a loop which reads in normalized frequencies
           (W) and ideal transfer function (Y)
      M=0
   30 M=M+1
      READ(5,21)W(M),Y(M)
   21 FORMAT(2F10.5)
C          The following writes the input data
      WRITE(6,22)M,W(M),Y(M)
   22 FORMAT(' I=',I3,' W=',D15.8,' V=',D15.8)
      IF(W(M).LT.1.D0)GOTO30
C          This DO loop gives value to the filter coefficients for the
           first iteration...the starting point
      DO 15 J=1,16
   15 X(J)=0.D
      X(4)=0.25
C          N is the number of coefficients - - - 4 for one section, 8 for
           two, etc.  LIMIT is the high bound on number of iterations.  The
           program will stop if more than LIMIT iterations are called.  EST
           and EPS are DFMFP parameters
   99 READ(5,60)N,LIMIT,EST,EPS
   60 FORMAT(2I5,2f10.5)
      WRITE(6,61)N,LIMIT,EST,EPS
   61 FORMAT('1N=',I3,' LIMIT=',I5,' EST=',D15.8,' EPS=',D15.8)
C          ICALL is the number of iterations (counting)
      ICALL=0
C          Conduct search for optimum coefficients
   98 CALL DFMFP(FUNCT,N,X,F,G,EST,EPS,25,IER,H)
C          Calculate roots for these coefficients
      CALL ROOTS(N,X)
C          IF calculated roots are outside the unit circle on the z-plane,
           invert them and set KFLAG...IF KFLAG=1, roots are outside the
           unit circle, if KFLAG=0, roots are inside
      CALL INSIDE(N,X,KFLAG)
      WRITE(6,26)IER,KFLAG,ICALL
```

```
   26  FORMAT('  IER=',I5,'  KFLAG=',I5,'  ICALL=',I5)
C         IER is the convergence flag
      IF((KFLAG,NE.0.0R.IER,NE.0).AND.ICALL.LE.LIMIT)GOTO98
      CALL ROOTS(N,X)
C         I CALL=-10 statement is to insure that final values will be printed
          out at the PRINT OUT stage in the FUNCT subroutine.
      ICALL=-10
      CALL FUNCT(N,X,F,G)
      END

      SUBROUTINE FUNCT(N,X,F,G)
C          This subroutine calculates the transfer function, the squared deviation
           from the ideal and the gradients.
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION H(184),X(16),G(16),YHT(100),E(100)
      COMPLEX*16 Z(100),NUM(100,4),DEN(100,4),Q,QBAR,ZCUR,ZCUR2
      COMMON/RAW/W(100),Y(100),M,ICALL
      PI=3.14159265358979
      K=N/4
      IF(ICALL.NE.0)GOTO101
      DO 102 I=1,M
  102 Z(I)=CDEXP(DCMPLX(0.DO,W(I)*PI)
C          If this is the first iteration (ICALL=0), then complex Z is
           defined.
C          Gain values A1 and A2 are initialized to zero
  101 A1=0.DO
      A2=0.DO
      DO 40 I=1,M
      ZCUR=Z(I)
      ZCUR2=ZCUR*ZCUR
      Q=DCMPLX(1,DO,0.DO)
      DO 33 J-1,K
      J4=(J-1)*4
      NUM(I,J)=1.DO+X(J4+1)*ZCUR+X(J4+2)*ZCUR2
      DEN(I,J)=1.DO+X(J4+3)*ZCUR+X(J4+4)*ZCUR2
   33 Q=Q*NUM(I,J)/DEN(I,J)
      QBAR=DCONJG(Q)
C           Calculate transfer function, numerator, denominator, and quotient
      YHT(I)=Q*QBAR
C           Calculate square of transfer function to get absolute value.
      A2=A2+YHT(I)
C           YHT is the calculated transfer function vaue
      YHT(I)=DSQRT(YHT(I))
   40 A1=A1+YHT(I)*Y(I)
      A=A1/A2
      F=0.DO
C         A is gain.  Insures that the transfer function is normalized for
          best fit to ideal.  F is functional value.  Accumulated squared
          deviation from ideal
      DO 57 J=1,16
   57 G(J)=0.DO
C         Calculation of F and gradient.  E is squared deviation for each
          individual frequency.  G is gradient
```

```
      QBAR=1.DO+X(J4+3)*ZCUR+X(J4+4)*ZCUR2
      A1=QBAR
      A2=(0.DO,-1.DO)*QBAR
      PHASE=PHASE-DATAN2(A2,A1)
   61 Q=Q/QBAR
      A1=Q*DCONJG(Q)
      A1=A*DSQRT(A1)
      PHASE=-1.DO*PHASE/PI
   60 WRITE(6,62)FREQ,PHASE,A1
   62 FORMAT('  W=',D15.8,' PHASE/PI=',D15.8,'  YHT=',D15.8)
      RETURN
      END
```

```
      DO 42 I=1,M
      ZCUR=Z(I)
      ZCUR2=ZCUR*ACUR
      YHT(I)=A*YHT(I)
      E(I)=YHT(I)-Y(I)
      F=F+E(I)**2
      DO 42 J=1,K
      J4=(J-1)*4
      Q=2.DO*E(I)*YHT(I)/NUM(I,J)
      G(J4+1)=G(J4+1)+Q*ZCUR
      G(J4+2)=G(J4+2)+Q*ZCUR2
      Q=-2.DO*E(I)*YHT(I)/DEN(I,J)
      G(J4+3)=G(J4+3)+Q*ZCUR
   42 G(J4+4)=G(J4+4)+Q*ZCUR2
      ICALL=ICALL+1
C         Increment iteration counter
C         Print values every 10th time
      IF((ICALL/10)*10.EQ.ICALL-1)WRITE96.25)ICALL,F
   25 FORMAT(' CALL NO.',I4,'  F=',D15.8)
      IF(ICALL.GT.0)RETURN
C         Print out stage.  For final pass, this is the printout section.
      WRITE(6,50)F
   50 FORMAT('  FINAL FUNCTION VALUE =',D15.8)
      WRITE(6,51)A
   51 FORMAT('  A=',D15.8)
      WRITE(6,52)(X(J),J=1,N)
   52 FORMAT(' FINAL X ='/(' ',4D15.8)
      WRITE(6,54)(G(J),J=1,N)
   54 FORMAT('  FINAL GRADIENT ='/('  ',4D15.8))
      DO 55 I=1,M
   55 WRITE96,56)I,W(I),Y(I),YHT(I),E(I)
   56 FORMAT(  I=',I3.'  W=',D15.8,' Y=;,D15.8,'  YHT=',D15.8,'  E=',D15.8)
      WRITE(6,59)N
   59 FORMAT(' FINAL TABLE FOR N =',I3)
C         Call Draw preparation.  OX is abscissa (frequency), OH is
          calculated transfer function values, OT is ideal transfer
          function values
      DO 60 I=1,201
      FREQ=.005DO*DFLOAT(I=1)
      ZCUR=CDEXP(DCMPLX(0.DO,FREQ*PI))
      ZCUR2=ACUR*SCUR
      Q=DCMPLX(1.DO,0.DO)
      PHASE=0.DO
      DO 61 J=1,K
      J4=(J-1)*4
C         Phase calculation.  Contributions from numerator and denominator
          are calculated separately.
      QBAR=1.DO+X(J4+1)*ZCUR+X(J4+2)*ZCUR2
      A1=QBAR
      A2=(0.DO,-1.DO)*QBAR
      PHASE=PHASE+DATAN2(A2,A1)
      Q=Q*QBAR
```

SUBROUTINE INSIDE: This subroutine checks to determine if discriminant is less than or equal(greater than) zero. Real roots and complex roots are handled separately.

```
      SUBROUTINE INSIDE(N,X,KFLAG)
      IMPLICIT REAL*8 (A-H,P-Z)
      DIMENSION X(16)
      J=-1
      KFLAG=0
   10 J=J+2
      IF(J.GT.N)RETURN
      B=-.5D0*X(J)
      C=X(J+1)
      DISC=B*B-C
      IF(DISC.LE.0.D0)GOTO20
C          Real Roots
      DISC=DSQRT(DISC)
      R1=B=DISC
      R2=B-DISC
      DR1=DABS(R1)
      DR2=DABS(R2)
      IF(DR1.LE.1.D0.AND.DR2.LE.1.D0)GOTO10
      KFLAG=1
      IF(DR1.GT.1.D0)R1=1.D0/R1
      IF(DR2.GT.1.D0)R2=1.D0/R2
      X(J)=-1.D0*(R1+R2)
      X(J+1)=R1*R2
      GOTO10
C           COMPLEX ROOTS
   20 IF(C.LE.1.D0)GOTO10
      KFLAG=1
      C=1.D0/C
      X(J+1)=C
      X(J)=X(J)*C
      GOTO10
      END
```

Subroutine Roots (N,X): Calculates roots to quadratics formed by calculated coefficients. Real roots and complex roots are handled separately.

```
      SUBROUTINE ROOTS(N,X)
      IMPLICIT REAL*8 (A-H,P-Z)
      DIMENSION X(16)
      WRITE(6,40)
   40 FORMAT(' ROOTS'/6X,'REAL',11X,'IMAG',11x,'REAL',11X,'IMAG')
      J=-1
   10 J=J+2
      IF(J.GT.N)RETURN
      B=-.5D0*X(J)
      C=X(J+1)
      DISC=B*B-C
      IF(DISC.LE.0.D0)GOTO20
C           Real roots
      DISC=DSQRT(DISC)
      R1=B+DISC
      R2=B-DISC
      WRITE(6,30)R1,R2
   30 FORMAT('  ',D15.8,15X,D15.8)
      GOTO10
C           Complex roots
   20 DISC=DSQRT(-1,D0*DISC)
      DISCM=-1.D0*DISC
      WRITE(6,50)B,DISC,B,DISCM
   50 FORMAT('  ',4D15.8)
      GOTO10
      END
```

LIST OF REFERENCES

(1) W. S. Boyle and G. E. Smith, "Charge-Coupled Semi-Conductor Devices", Bell Sys. Tech. Journal. 49: 487, 1970.

(2) G. F. Amelio, W. J. Bertram, and M. F. Tompsett, "Charge-Coupled Imaging Devices: Design Considerations", IEEE Transactions on Electron Devices, Vol. ED-18, No. 11, Nov. 1971.

(3) M. F. Tompsett, "Charge-Coupled Imaging Devices: Experimental Results", IEEE Transactions on Electron Devices, vol. ED-18, No. 11, Nov. 1971.

(4) F. L. J. Sangster, "Integrated MOS and Bipolar Analog Delay Lines Using Bucket Brigade Capacitor Storage", 1970 IEEE Solid State Circuits Conference, Digital Tech. Papers, pp. 74-75, 185, 1970.

(5) D. R. Collins , W. H. Bailey, W. M. Gosney, and D. D. Buss, "Charge-Coupled Device Analog Matched Filters", Electronic Letters, 8, pp. 328-329, June 1972.

(6) J. E. Bounden, R. Eames, and J. B. Roberts, Proceedings of the International Conference on the Technology and Applications of CCD, Edinburgh, Scotland, p. 206, Sept., 1974.

(7) T. F. Tao, V. Iamsa-ad, S. V. Holmes, B. Freund, L. Saetre, and T. R. Zimmerman, "Sampled Analog CCD Recursive Comb Filters", Proceedings of the Third International Conference on Applications of Charge Coupled Devices, 1975.

(8) C. H. Sequin and A. M. Mohsen, "Linearity of Electrical Charge Injection into Charge-Coupled Devices", IEEE Journal of Solid State Circuits, May 1975.

(9) M. I. Skolnik, Introduction to Radar Systems, 1962.

(10) W. J. Butler, W. E. Engeler, H. S. Goldberg, C. M. Puckette, H. Lobenstein, "Charge Transfer Analog Memories for Radar Systems", IEEE Advanced Solid-State Components for Signal Processing, IEEE Catalog no. 75CH0979-5CAS, 1975.

(11) B. C. (Tsung-i) Kuo, Automatic Control Systems, Prentice Hall, 1962.

(12) B. C. (Tsung-i) Kuo, Analysis and Synthesis of Sampled Data Control Systems, Prentice Hall, 1963.

(13) A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Prentice Hall, 1975.

(14) J. L. Herrero and G. Willoner, Synthesis of Filters, Prentice Hall, 1966.

(15) R. Wolfe (editor), Applied Solid State Science, "Advances in Materials and Device Research", Academic Press, Inc., 1969.

(16) K. Steiglitz, "Computer-Aided Design of Recursive Digital Filters", IEEE Transactions on Audio and Electroacoustics, vol. AU-18, No. 2, June 1970.

(17) W. D. Baker and D. F. Barbe, "Charge-Coupled Device Applications", Naval Research Reviews, Dec. 1974.

(18) D. A. Smith, W. J. Butler, C. M. Puckette, "Programmable Bandpass Filter and Tone Generator Using Bucket Brigade Delay Lines", IEEE Transactions on Communications, vol. COMM-20, No. 7, July 1974.

(19) Y. Daimon, A. M. Mohsen, and T. C. McGill, "Final Stage of the Charge Transfer Process in Charge Coupled Devices", IEEE Transactions on Electron Devices, Vol. ED-21, No. 4, April 1974.

(20) A. M. Mohsen, T. C. McGill, and C. A. Mead, "Charge Transfer in Overlapping Gate Charge-Coupled Devices", IEEE Journal of Solid State Circuits, vol. SC-8, No. 3, June 1973.

(21) W. H. Chang and L. G. Heller, "Structure Dependence of Free Charge Transfer in Charge-Coupled Devices", IBM Journal of Research Development, Sept. 1974.

(22) D. F. Barbe, "The Charge-Coupled Concept", Report of NRL Progress, March 1972.

(23) K. H. Zaininger, et. al. "Charge Coupled Devices", Phase I, Final Report, RCA Laboratories, David Sarnoff Research Center. NAVELEX contract N00039-73-C-0014, March 1973.

(24) "Charge Coupled Devices in Signal Processing Systems", Vol. II. Analog Signal Processing. TRW Systems Group. NAVELEX/NRL contract N00014-74-C-0068, December 1974.

(25) M. H. White and W. R. Webb, "Study of the Use of Charge-Coupled Devices in Analog Signal Processing Systems", Final Report, Westinghouse Defense and Electronic Systems Center. NRL contract No. N00014-74-C-0069, May 1974.

(26) Bernard Gold and C. M. Rader, Digital Processing of Signals, McGraw-Hill, 1969.

(27) L. T. Saetre, "Sampled Analog Recursive Comb Filters and Their Application to MTI Radar", Naval Postgraduate School Master's Thesis, December 1975.

(28) Veerachai Iamsa-ad, "Charge-Coupled Devices for Analog Signal Processing - Recursive Filter Study", Naval Postgraduate School Engineer's Thesis, December 1974.

(29) Bruce R. Freund, "Implementation of Comb Filters by Sampled Analog Techniques", Naval Postgraduate School Master's Thesis, June 1975.

(30) L. R. Rabiner and Bernard Gold, Theory and Application of Digital Signal Processing, Prentice Hall, 1975.

(31) Herbert Taub and J. R. Schilling, Principles of Communications Engineering, Prentice Hall, 1972.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 022 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 52 Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	2
4.	Assoc. Professor T. F. Tao, Code 52 TV Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	5
5.	Ronald Stoutmeyer Code 3921 Naval Weapons Center China Lake, California 93555	1
6.	LCDR Stacy V. Holmes, USN Office of Naval Research Arlington, Virginia 20390	2
7.	LT Ang Vong Mongkol Royal Cambodian Navy P. O. Box BF Pacific Grove, California 93940	1